

New constraint on cosmological variation of the proton-to-electron mass ratio from Q0528–250

Julian A. King^{1,3*}, Michael T. Murphy^{2†}, Wim Ubachs^{3‡}, John K. Webb^{1§}.

¹*School of Physics, University of New South Wales, Sydney, NSW, 2052, Australia*

²*Centre for Astrophysics and Supercomputing, Swinburne University of Technology, Victoria, 3122, Australia*

³*Department of Physics and Astronomy, LaserLaB, VU University, De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands*

Accepted for publication, 23 June 2011

ABSTRACT

Molecular hydrogen transitions in quasar spectra can be used to constrain variation in the proton-to-electron mass ratio, $\mu \equiv m_p/m_e$, at high redshifts ($z \gtrsim 2$). We present here an analysis of a new spectrum of the quasar Q0528–250, obtained on VLT/UVES (the Ultraviolet and Visual Echelle Spectrograph, on the Very Large Telescope), and analyse the well-known H₂ absorber at $z = 2.811$ in this spectrum. For the first time we detect HD (deuterated molecular hydrogen) in this system with a column density of $\log_{10}(N/\text{cm}^{-2}) = 13.27 \pm 0.07$; HD is sensitive to variation in μ , and so we include it in our analysis. Using 76 H₂ and 7 HD transitions we constrain variation in μ from the current laboratory value to be $\Delta\mu/\mu = (0.3 \pm 3.2_{\text{stat}} \pm 1.9_{\text{sys}}) \times 10^{-6}$, which is consistent with no cosmological variation in μ , as well as with previous results from other H₂/HD absorbers. The main sources of systematic uncertainty relate to accurate wavelength calibration of the spectra and the re-dispersion of multiple telescope exposures onto the one pixel grid.

Key words: quasars: absorption lines, cosmology: observations, methods: data analysis

1 INTRODUCTION

The Standard Model of particle physics contains a number of “fundamental constants”, the values of which are predicted by the theory, and which must be measured by experiment. Two dimensionless constants of particular importance are the fine-structure constant, $\alpha \equiv e^2/(4\pi\epsilon_0\hbar c)$, and the proton-to-electron mass ratio, $\mu \equiv m_p/m_e$. The former determines the strength of electromagnetism, whereas the latter is a measure of the relative strengths of the strong and electroweak scales. Detection of spatial or temporal variation in a fundamental constant would directly demonstrate the incompleteness of the Standard Model, and might guide efforts to develop post-Standard Model theories.

Quasar absorption lines can be used to constrain variations in α and μ across most of the observable universe. In particular, the wavelengths of certain transitions are sensitive to changes in the value of α or μ ; if α or μ were different, the relative wavelengths of the spectral lines would differ from those observed in the laboratory, even after correction for the redshift of the absorbing cloud. By comparing measurements of absorption transitions in quasar absorbers to precise laboratory measurements, one is able to constrain changes in α or μ .

1.1 Molecular hydrogen

Thompson (1975) noted that the Lyman and Werner transitions of molecular hydrogen (with $\lambda_0 \lesssim 1150\text{\AA}$) can be used to constrain variations in μ at $z \gtrsim 2$, where these UV transitions are redshifted into the optical spectrum and therefore can be observed with ground based telescopes. In particular, one can measure the quantity $\Delta\mu/\mu \equiv (\mu_z - \mu_0)/\mu_0$ where μ_z is the measurement of μ at a redshift z and μ_0 is the laboratory value of μ .

If μ were to vary, the wavelengths of the H₂ transitions would vary by different amounts. The shift can be quantified as

$$\lambda_i = \lambda_i^0 (1+z) \left(1 + K_i \frac{\Delta\mu}{\mu} \right), \quad (1)$$

where λ_i^0 is the rest wavelength of the transition, λ_i is the observed wavelength and K_i is a ‘sensitivity coefficient’ determining the magnitude and sign of the effect (Varsholovich & Levshakov 1993). That is,

$$K_i = \frac{d \ln \lambda_i}{d \ln \mu}. \quad (2)$$

z is the absorption redshift of the cloud (equivalently, the redshift of a transition with no sensitivity to a change in μ). The determination of z is not degenerate with determination of $\Delta\mu/\mu$ provided that two or more transitions with different K_i values are used.

It is worth noting that, if we define a nominal observation red-

* E-mail: jking.phys@gmail.com

† E-mail: mmurphy@swin.edu.au

‡ E-mail: w.m.g.ubachs@vu.nl

§ E-mail: jkw@phys.unsw.edu.au

shift for a particular transition as z_i , then

$$\frac{\Delta v_i}{c} \approx \zeta_i \equiv \frac{z_i - z}{1 + z} = K_i \frac{\Delta\mu}{\mu}. \quad (3)$$

That is, the velocity shift of an individual line is linearly proportional to $\Delta\mu/\mu$. ζ_i is known as the ‘reduced redshift’. The redshift z is that of a transition with $K_i = 0$. In practice, z must be measured simultaneously with ζ_i from a set of transitions (e.g. through regression methods) because it is unknown *a priori*.

Recent indications have emerged for a significantly positive $\Delta\mu/\mu$ from the study of two H₂ absorbers at $z = 2.595$ and $z = 3.025$ toward Q0405–443 and Q0347–383 respectively. Ivanchik et al. (2005) analysed high resolving power ($R \sim 50,000$) VLT/UVES spectra of these absorbers, and reported a combined value of $\Delta\mu/\mu$ for the two absorbers of $(30.5 \pm 7.5) \times 10^{-6}$ and $(16.5 \pm 7.4) \times 10^{-6}$ depending on whether H₂ laboratory wavelengths were used from Abgrall et al. (1993) or Philip et al. (2004) respectively. Reinhold et al. (2006) analysed the same two absorbers using improved laboratory wavelength values, and K_i values derived from a more accurate calculation. Using a table of observed line wavelengths from Ivanchik et al. (2005), they derived $\Delta\mu/\mu$ values of $(27.8 \pm 8.8) \times 10^{-6}$ for the absorber toward Q0405–443, $(20.6 \pm 7.9) \times 10^{-6}$ for the absorber toward Q0347–383, and a combined value of $(24.5 \pm 5.9) \times 10^{-6}$.

King et al. (2008) analysed the same raw spectra of Q0405–443 and Q0347–383 using a more comprehensive fitting method in which all transitions are fitted simultaneously with the addition of $\Delta\mu/\mu$ as a free parameter in the fit. They used an improved flux extraction and, importantly, an improved wavelength calibration procedure (Murphy et al. 2007) which should significantly reduce systematic errors. They found $\Delta\mu/\mu = (8.2 \pm 7.4) \times 10^{-6}$ and $(10.1 \pm 6.2) \times 10^{-6}$ for the absorbers in Q0347–383 and Q0405–443 respectively. They also analysed the $z = 2.811$ absorber toward Q0528–250, providing a constraint of $\Delta\mu/\mu = (-1.4 \pm 3.9) \times 10^{-6}$. A weighted mean of these three values yielded $\Delta\mu/\mu = (2.6 \pm 3.0) \times 10^{-6}$.

Wendt & Reimers (2008) and Thompson et al. (2009) also analysed the absorbers toward Q0405–443 and Q0347–383 using the same UVES spectra, and derived results statistically consistent with $\Delta\mu/\mu = 0$. Wendt & Molaro (2011) re-analysed the existing Q0347–383 exposures combined with exposures under program ID 68.B-0115(A) that had not been previously analysed, again deriving a result statistically consistent with zero. Given the results of King et al. (2008), Wendt & Reimers (2008), Thompson et al. (2009) and Wendt & Molaro (2011) it does not appear that there is significant evidence for $\Delta\mu/\mu \neq 0$ from the absorbers in Q0405–443 and Q0347–383.

Malec et al. (2010) reported a strong constraint from the $z = 2.059$ absorber toward J1213–0050 from spectra obtained using Keck/HIRES. A combination of good seeing conditions and the brightness of the quasar (r -band magnitude ≈ 16.5 mag) meant that they were able to obtain a $R \sim 110,000$ spectrum of sufficiently high SNR in a single night. They found that $\Delta\mu/\mu = (5.6 \pm 5.5_{\text{stat}} \pm 2.9_{\text{sys}}) \times 10^{-6}$ from 86 H₂ transitions. Recently the same J1213–0050 absorber system was observed with VLT/UVES at a somewhat lower resolution ($R \sim 53,000$) but with a higher signal-to-noise ratio (van Weerdenburg et al. 2011). The spectrum covered 96 H₂ and HD lines and yields a result of $\Delta\mu/\mu = (8.5 \pm 3.6_{\text{stat}} \pm 2.2_{\text{sys}}) \times 10^{-6}$, in very good agreement with the results from Keck.

From consideration of the results above, it does not appear at

present that there is any strong evidence for $\Delta\mu/\mu \neq 0$ at $z > 1$ from measurements of H₂ absorbers.

Unfortunately the number of quasars known to contain H₂ absorption features in their spectra is low ($\lesssim 15$), which makes obtaining a statistical sample of $\Delta\mu/\mu$ measures from the H₂ method difficult. To obtain a reliable measurement of $\Delta\mu/\mu$ from H₂, one needs a H₂ column density that is high enough to produce significant absorption, but preferably not saturated (which makes determination of line centroids difficult). Similarly, transitions of significantly different K_i must be used to achieve good precision on the measurement of $\Delta\mu/\mu$. Unfortunately, the only quasar absorbers to date which have been able to yield measurements of $\Delta\mu/\mu$ at the $\lesssim 10^{-5}$ level are those listed above, which explains the significant number of analyses on these few objects.

1.2 Historical Q0528–250 constraints

Since the absorber in Q0528–250 is the focus of this work, we note here previous constraints on variation in μ from this object.

Varshalovich & Levshakov (1993) analysed the $R \sim 4,000$ spectrum of Foltz et al. (1988) of Q0528–250 to obtain $|\Delta\mu/\mu| < 0.005$. Varshalovich & Potekhin (1995) reanalysed the same spectrum to conclude that $|\Delta\mu/\mu| < 0.002$. Potekhin et al. (1998) used new observations of the same system at higher resolving power ($R \sim 14,000$) to obtain $\Delta\mu/\mu = (-10 \pm 8) \times 10^{-5}$.

Cowie & Songaila (1995) used a $R = 36,000$ Keck observation of Q0528–250 to produce $\Delta\mu/\mu \in [-7, 5.5] \times 10^{-4}$ (95 percent confidence limits).

Ubachs & Reinhold (2004) used the improved laboratory H₂ wavelengths of Philip et al. (2004) to re-investigate the reported line positions for transitions in the spectra of Q0528–250, Q0347–383 and Q1232+082. For the combined data, they found that $\Delta\mu/\mu = (-0.5 \pm 1.8) \times 10^{-5}$. Omitting the Q0528–250 data, which was of poorer quality, they obtained $\Delta\mu/\mu = (1.9 \pm 1.5) \times 10^{-5}$. An explicit value for Q0528–250 was not given.

1.3 Other quasar constraints on $\Delta\mu/\mu$

The inversion transitions of ammonia are strongly sensitive to variation in μ , with $K_i \sim 4.2$ (Flambaum & Kozlov 2007). Murphy et al. (2008a) and Henkel et al. (2009) compared the inversion transitions of NH₃ with rotational molecules to obtain stringent constraints on $\Delta\mu/\mu$ at $z < 1$. Murphy et al. (2008a) used B0218+357 to obtain $\Delta\mu/\mu = (0.74 \pm 0.47_{\text{stat}} \pm 0.76_{\text{sys}}) \times 10^{-6}$ at $z = 0.69$, whilst Henkel et al. (2009) used PKS1830–211 to obtain $\Delta\mu/\mu = (0.08 \pm 0.47_{\text{sys}}) \times 10^{-6}$ at $z = 0.89$. Kanekar (2011) analysed the $z = 0.69$ absorber toward B0218+357 to obtain $\Delta\mu/\mu = (-3.5 \pm 1.2) \times 10^{-7}$.

Constraints derived from the analysis of H₂/HD absorbers are not statistically competitive with the ammonia results at present. However, the ammonia results are from lower redshifts ($z < 1$), whilst the H₂/HD results all probe $z > 2$, and therefore probe much larger distances and lookback times. The two methods are thus complementary.

Recently it has been suggested that H₃O⁺ ions (Kozlov & Levshakov 2011) and CH₃OH molecules (Jansen et al. 2011) are even more sensitive probes to detect or constrain μ variation; these transitions exhibit large sensitivity coefficients ($K_i > 10$).

1.4 Motivation for re-analysis of Q0528–250

The constraint on $\Delta\mu/\mu$ derived from the Q0528–250 absorber presented in King et al. (2008) is very precise. However, there are good reasons to revisit this absorber. Some of the exposures which contributed to the spectrum used in that analysis are not well-calibrated. By well-calibrated, we mean that the exposures do not have ThAr (thorium-argon) calibration spectra taken immediately afterwards. The design of VLT/UVES is such that the position of the spectrograph grating is reset between different quasar exposures. Although the specification is such that the placement of the grating should be good to within 0.1 pixels (D’Odorico et al. 2000), the use of spectra for which the ThAr spectra were taken after grating resets necessarily introduces wavelength calibration uncertainties into the spectrum.

A more subtle concern for the earlier observations (under program IDs P66.A-0594, P68.A-0600 and P68.A-0106) is that they suffer from the fact that the slit width used for the observations was often significantly larger than the prevailing seeing conditions. For those observations, a 1 arcsecond slit was used in both the blue and red arms of UVES. The average ratio of the slit width to seeing, where seeing is quantified by the output of the DIMM (differential image motion monitor) and the seeing values were weighted by the duration of the exposure, was 0.78. In one exposure of 1.6 hours, the ratio was as low as 0.48. In the case where the seeing is significantly smaller than the slit width, the instrumental profile will be non-Gaussian, which complicates the analysis. Although in principle one can use a numerically-provided instrumental profile to convolve the Voigt profile model with, assuming that the instrumental profile is Gaussian is significantly easier.

2 DATA

2.1 Description of the new Q0528–250 spectrum

Note that this paper uses only new spectra of Q0528–250 so as to measure $\Delta\mu/\mu$ as independently as possible from that presented in King et al. (2008).

Our new observations are summarized in Table 1. We observed Q0528–250 in late 2008/early 2009 under ESO program ID 082.A-0087, with exposures totalling approximately 8.1 hours. All science exposures were followed immediately by ThAr calibration exposures without any intervening grating resets, thereby reducing drifts and shifts in the spectrograph in the intervening time. The average slit-width-to-seeing ratio for these exposures is 1.03; that is, the slit width was more appropriately matched to the average seeing conditions than for the exposures contributing to the spectrum analysed in King et al. (2008).

We optimally extracted and wavelength calibrated the raw spectra using the ESO UVES Common Pipeline Language (CPL) software suite. While the CPL code redisperses the spectra onto a linear wavelength scale by default, we used only the original, ‘unredispersed’ flux and error array for each echelle order in subsequent reduction steps. A custom code, UVES_POPLER¹, was used to combine these extracted echelle orders from the 10 exposures into a single spectrum. The relevant heliocentric corrections and air–vacuum wavelength conversions were calculated and applied to the wavelength scale of each echelle order before redispersing

all orders onto a common, log-linear wavelength scale with dispersion $2.0 \text{ km s}^{-1} \text{ pixel}^{-1}$. The best-fitting relative flux scaling between overlapping orders from all exposures was determined using an automatic χ^2 minimisation scheme. After applying these relative scalings, the extracted quasar exposures were coadded with inverse-variance weighting and a cosmic ray rejection algorithm. Note that, as implied by equation 1, effects which shift the wavelength scale by a constant velocity at all wavelengths are generally unimportant².

UVES_POPLER was also used to automatically fit a continuum to the final spectrum. This continuum was acceptable for our purposes only redwards of the Lyman- α forest. In the Lyman- α forest the continuum was manually re-fitted with low-order polynomials to obtain a nominal continuum against which the H_2 absorption lines of interest could be reliably defined. Local constant or linear modifications to this continuum were necessary in the vicinity of some H_2 lines when conducting the detailed spectral fitting, as described below.

Using the CPL pipeline we extracted the ThAr flux with the same spatial profile weights derived for the corresponding quasar exposure, thereby allowing the wavelength calibration polynomial most appropriate to the quasar exposure to be established. The CPL pipeline makes use of the ThAr line-list derived by Murphy et al. (2007) specifically for accurate calibration of UVES ThAr exposures. The calibration residuals had a root-mean-square (RMS) deviation around the final polynomial of 70 m s^{-1} in UVES’s blue arm (where all the H_2 lines fall) and 55 m s^{-1} over the majority of the two red arm chips, on average.

From the widths of the extracted ThAr features we derived an average full-width-at-half-maximum (FWHM) resolution of $\approx 5.45 \text{ km s}^{-1}$, as expected for a slit width of $0''.8$ (D’Odorico et al. 2000). We initially used this FWHM value as the instrumental resolution in our spectral fits. However, we expect – and later demonstrate – that a somewhat lower value of 5.15 km s^{-1} better represents narrow, unresolved lines in the spectrum because the quasar illuminates the slit centrally rather than uniformly like the ThAr lamp.

2.2 Laboratory wavelengths and K_i values

Until recently, the uncertainties in the laboratory wavelengths for the Lyman and Werner series H_2 transitions were comparable to the typical uncertainties in the measured line positions in the quasar spectra. However, a significant amount of recent work has improved this situation by using laser-based spectroscopic techniques to obtain much more precise laboratory values. In particular, Philip et al. (2004), Ubachs & Reinhold (2004) and Ivanov et al. (2008b) improved the fractional wavelength accuracy ($\delta\lambda/\lambda$) to $\sim 5 \times 10^{-8}$, whilst Salumbides et al. (2008) improved the fractional accuracy to $\sim 5 \times 10^{-9}$ for most Lyman transitions and $\sim 10^{-8}$ for Werner transitions. See also Bailly et al. (2009).

The K_i coefficients can be calculated either through semi-empirical methods (Ubachs et al. 2007) or from *ab initio* methods (Meshkov et al. 2006). The K_i values derived from these two processes are in good agreement (to within 1 percent).

¹ See http://astronomy.swin.edu.au/~mmurphy/UVES_popler; maintained by MTM.

² This is not strictly true when many quasar exposures are combined, as is the case here. If different velocity shifts are applied to the different quasar exposures, and if the relative weights of the exposures vary with wavelength when forming the final, combined spectrum, then small relative velocity shifts will be measured between transitions at different wavelengths.

Table 1. Journal of VLT/UVES observations of quasar Q0528–250. Each row specifies the starting time of a single quasar exposure (column 1), all of which were taken with the DIC1 dichroic splitter, giving a central wavelength (λ_c) in the blue (“B”) and red (“R”) arms as shown in column 2. The third column gives the exposure times for the quasar and subsequent ThAr wavelength calibration exposures in each arm. Column 4 gives the slit width used for each arm. These should be compared with the average FWHM seeing, as reported by the on-site Differential Image Motion Monitor (DIMM). The average difference between the spectrograph air temperatures at the times of the quasar and ThAr exposures, $\Delta T \equiv T_{\text{QSO}} - T_{\text{ThAr}}$, is given in the sixth column. The final column shows the atmospheric pressure difference between the quasar and ThAr exposures, $\Delta P \equiv P_{\text{QSO}} - P_{\text{ThAr}}$. On-chip binning by a factor of 2 in both spatial and spectral directions was used for all exposures.

UT date & time [yyyy-mm-dd hh:mm]	λ_c [nm]		T_{exp} [s]			Slit-width [arcseconds]		DIMM seeing [arcseconds]	ΔT [K]		ΔP [hPa]
	B	R	QSO	B	R	B	R		B	R	
2008-11-23 03:58	390	564	2900.0	1.6	1.8	0.8	0.7	0.75	0.2	0.0	0.2
2008-11-23 04:52	390	580	2900.0	1.6	1.8	0.8	0.7	0.79	0.0	0.0	0.2
2008-11-25 07:35	390	580	2900.0	1.6	1.8	0.8	0.7	0.43	0.0	0.0	0.0
2008-12-23 05:03	390	580	2900.0	1.6	1.8	0.8	0.7	0.75	0.0	0.0	0.0
2008-12-23 05:54	390	580	2900.0	1.6	1.8	0.8	0.7	0.74	0.0	0.0	0.1
2008-12-23 06:46	390	580	2900.0	1.6	1.8	0.8	0.7	0.99	0.0	0.0	0.2
2008-12-23 07:38	390	580	2900.0	1.6	1.8	0.8	0.7	1.07	0.1	0.0	0.1
2009-01-25 02:40	390	580	2900.0	1.6	1.8	0.8	0.7	0.72	0.0	0.1	0.0
2009-01-26 04:02	390	564	2900.0	1.6	1.8	0.8	0.7	1.19	0.0	0.2	0.1
2009-02-26 02:52	390	580	2900.0	1.6	1.8	0.8	0.7	0.75	0.0	0.0	0.0

Malec et al. (2010) (see their Table 1) have collated the laboratory wavelength data from the above sources and the K_i values from Ubachs et al. (2007) to produce a comprehensive list which may be used for an analysis of $\Delta\mu/\mu$. They have used oscillator strengths from Abgrall et al. (1994) and calculated damping coefficients from Abgrall et al. (2000). We therefore have used the laboratory wavelength data, K_i values, oscillator strengths and damping coefficients tabulated in Malec et al. (2010) for our analysis of the H₂ absorber in Q0528–250.

2.2.1 HD

We have detected HD (deuterated molecular hydrogen) at $z \sim 2.811$ in the Q0528–250 spectrum with a column density of $\log_{10}(N/\text{cm}^{-2}) = 13.27 \pm 0.08$. Malec et al. (2010) have collated oscillator strengths, laboratory wavelength values and K_i values for HD; wavelength values are from Hollenstein et al. (2006) and Ivanov et al. (2008a), K_i values are from Ivanov et al. (2008a, 2010) and oscillator strengths were calculated by Malec et al. (2010) from Einstein A coefficients given in Abgrall & Roueff (2006). We describe our treatment of HD further in section 3.5.

3 METHODS AND METHODOLOGY

To estimate $\Delta\mu/\mu$ in Q0528–250, we model the H₂/HD transitions in the absorber with Voigt profiles using the non-linear least-squares χ^2 minimisation program VPFIT³. VPFIT was written specifically to fit Voigt profiles to quasar absorption spectra. Each Voigt profile is described by three parameters: the redshift of the transition, z , the column density, N , and the Doppler width, b . Knowledge is also required of the rest wavelength, oscillator strength (f) and damping constant (Γ) for each transition. The total model is the sum of the optical depths of a series of Voigt profiles convolved with a model for the instrumental profile (assumed to be Gaussian).

If each transition in the absorber was well-represented by a single, unblended Voigt profile then determining $\Delta\mu/\mu$ would be

straightforward: one would measure a redshift for each H₂ transition, and then calculate a set of ζ_i values (equation 3). Equation 3 implies that a plot of ζ_i against K_i would have a slope of $\Delta\mu/\mu$.

In practice, there are two factors which complicate the analysis. Firstly, the H₂ transitions all have rest wavelengths less than that of Lyman- α (1215.7Å), and therefore are found only in the Lyman- α forest, the dense series of broad H I transitions found bluewards of the quasar’s Lyman- α emission line. This means that the H₂ transitions are seen against a background of relatively broad Lyman- α lines found at random redshifts and with varying optical depths. To accurately determine the line centroids of the H₂ transitions, one must model the Lyman- α forest transitions simultaneously with the H₂ transitions. Similarly, unless one allows the parameters of the Lyman- α transitions to be determined simultaneously with those of the H₂ transitions, one will necessarily underestimate the errors in the line centroids for the H₂ transitions and therefore also the error on $\Delta\mu/\mu$.

Secondly, the Q0528–250 absorber is not well-represented by a single Voigt profile. Instead, it shows ‘velocity structure’ – several clouds closely separated in velocity space but having different optical depths and Doppler widths. Provided that the SNR of the spectrum is sufficiently high, one can decompose the observed absorption profile into different ‘velocity components’ (one for each absorption cloud). A reasonable model of the absorber is necessary if one is to obtain an accurate estimate of $\Delta\mu/\mu$.

In order to estimate $\Delta\mu/\mu$ for the Q0528–250 absorber it is therefore necessary to: *i*) construct a Voigt profile model for the Q0528–250 H₂ absorber, determining the required number of velocity components using objective criteria, and; *ii*) construct a model for the Lyman- α forest in the vicinity of each of the H₂ transitions. These two steps constitute almost all of the effort in accurately measuring $\Delta\mu/\mu$ from an existing spectrum of the absorber (significant work being required to obtain a spectrum). In section 3.1 we discuss the statistical criteria used to determine the preferred model. In section 3.2 we discuss how we build up a model of the Lyman- α forest in the vicinity of the H₂ transitions. In section 3.2.5 we describe how we determine the best-fitting number of H₂ velocity components.

³ <http://www.ast.cam.ac.uk/~rfc/vpfit.html>

3.1 Model selection

Webb et al. (1999, 2001, 2011), Murphy et al. (2001, 2003, 2004, 2008b) and King et al. (2011) discussed the importance of accurately constructing a Voigt profile fit to an absorber in the context of potential variation of α . We apply a similar methodology here in attempting to create the simplest Voigt profile model which best explains the absorption spectrum. King et al. (2008) and Malec et al. (2010) discussed the importance of using an appropriate number of velocity components when fitting H_2 absorbers to estimate $\Delta\mu/\mu$.

When comparing two models with the same number of free parameters, under the maximum likelihood method one should prefer whichever model has the lowest χ^2 . However, when comparing models with different numbers of free parameters, one must consider whether the reduction in χ^2 as a result of the addition of more free parameters is sufficiently large to justify the extra free parameters; a so-called ‘information criterion’ should be used to select the most appropriate model.

One statistic which can be used to compare models with different numbers of free parameters is the Akaike Information Criterion (Akaike 1974), defined as $\text{AIC} = \chi^2 + 2p$, where p is the number of free parameters. As the AIC is only correct in the limit of large n/p (where n is the number of spectral points included in the fit), we use the AIC corrected for finite sample sizes, (Sugiura 1978), defined as

$$\text{AICC} = \chi^2 + 2p + \frac{2p(p+1)}{n-p-1}. \quad (4)$$

The number of degrees of freedom, ν , is therefore given by $\nu = n - p$. If several competing models are being considered, one chooses the model which has the lowest AICC. The actual value of the AICC is not important; only relative differences matter. The suggested interpretation scale for the AICC is the Jeffreys’ scale (Jeffreys 1961) (see Liddle 2007, for further discussion), where $\Delta\text{AICC} = 5$ and $\Delta\text{AICC} = 10$ are considered strong and very strong evidence against the weaker model respectively.

When constructing the Voigt profile model, our goal is to find a physically plausible model for the spectral data which minimises the AICC.

3.2 Constructing the Voigt profile model

We built up our model of the molecular hydrogen transitions and the surrounding forest through an iterative process which we describe here. With knowledge of the redshift of the molecular hydrogen absorbers, in each spectrum we searched for molecular hydrogen transitions which we considered to be potentially usable. We consider potentially usable transitions to be those for which the molecular hydrogen transition can be visually distinguished from its surrounds. This necessarily precludes the use of H_2 transitions in regions of near zero flux, but in any event these transitions would contribute no meaningful constraint on $\Delta\mu/\mu$.

3.2.1 Model construction process and physical assumptions

We did not build up our region fits from those used in King et al. (2008), but rather started afresh in order to provide as independent a measurement of $\Delta\mu/\mu$ as possible. From a list of potentially usable transitions, we then selected a spectral region around the H_2 transition, where the region should be large enough to include any absorption feature which might overlap with the H_2 transition. In

general, we attempted to ensure that the fitting region was sufficiently large so as to return to the local continuum, although this was not always possible. In each of the fitting regions, we modelled the molecular hydrogen transition and then modelled all surrounding features as H I. To do this, we added and removed H I components to attempt to achieve a statistically satisfactory model, using the criterion set out in section 3.1. The line parameters (N, b, z) for different H I transitions are independent of each other. Note that although most transitions observed in the forest are indeed due to H I, there are also metal transitions from other absorbers along the line of sight (including galactic and atmospheric lines). The identification of the origin of these transitions is not necessary if they do not overlap with the H_2 transitions; we simply modelled them as H I in order to have a physical model for them. We describe the treatment of metal lines which overlap with H_2 lines below. For all transitions assumed to be H I (which we refer to hereafter as just H I transitions), we use only the $\lambda 1215.7$ transition rather than the whole Lyman series, to prevent line misidentification spuriously impacting regions blueward of that transition. Where Lyman- β transitions exist in the blue region of the spectrum, we simply modelled them with additional Lyman- α components.

We then combined models from the regions fitted individually into a model where the regions are fitted simultaneously. As the line parameters for the individual H_2 transitions were independent when the regions were fitted independently, at this stage we imposed physical restrictions on the transitions by tying certain parameters together to reflect the fact that the transitions are physically connected. The ground states of the H_2 and HD molecules have several rotational sub-states described by the quantum number J . We refer to these sub-states as being different J -levels. For H_2 and HD each J -level has a different ground state population. This means that the column density, N , should be the same for each velocity component arising from a particular J -level. The same applies for the Doppler width, b , and the redshift, z . We impose the assumption that the velocity structure is the same in all the J -levels, that is that a velocity component has the same value of z in all J -levels. We describe later the impact of this assumption.

King et al. (2008) did not actually impose the requirement that each velocity component arising from the same J -level has the same N ; instead they fitted the column densities for each transition as free parameters. We apply a similar approach. We must, however, ensure that a physical consistency is maintained, in that the ratios of the line strengths between different velocity components should be the same for transitions arising from the same J -level. We therefore imposed the requirement that the ratio of the column densities between the different components was the same for transitions arising from the same J -level. In this way, the total column density (effectively, oscillator strength) for each transition was a free parameter, but the ratios of the individual column densities of different velocity components within each transition were constrained.

To model the forest, we thus iteratively refined the fit by alternately allowing VPFIT to minimise χ^2 for a particular model, then attempted to improve that model through the addition and deletion of H I components to obtain a robust model according to the criteria in section 3.1.

3.2.2 Blending with metal lines

During the iterative process, it can become clear that a molecular transition is blended with another line (presumed H I) when it was not thought to be from a fit to just that region. This is because

the information from the other molecular hydrogen transitions imposes a strong constraint on the b parameters and redshifts of that transition, thus uncovering apparently hidden blends. These blends necessitate the addition of H I components that overlap with the H₂ or HD transition in question. With the addition of extra H I transitions, an acceptable fit can generally be achieved. This demonstrates the utility of fitting all transitions simultaneously; otherwise-inconspicuous blends are generally revealed.

In a few instances, the transitions which had to be included to achieve a statistically acceptable fit had extremely narrow b parameters ($b \lesssim 5 \text{ km s}^{-1}$). In this case, it is likely that the blend is a metal line from an unknown absorber along the line of sight. As a result, we excluded such H₂ or HD transitions from our fit. The reason for not accepting transitions affected by narrow- b interlopers is that any inaccuracy in modelling the interloping transitions could lead to a significant bias in measuring the H₂ line position – the narrow b parameter(s) of the interloping transition(s) means that the absorption they cause varies rapidly across the H₂ line profile. Ultimately, the joint fit of all the molecular hydrogen transitions allows the detection and rejection of transitions which are likely to be contaminated by metal lines.

Rejecting transitions which are suspected to be contaminated cannot bias $\Delta\mu/\mu$ away from zero. Moreover, this should not bias $\Delta\mu/\mu$ significantly. If the suspicion of contamination in particular lines was in fact due to $\Delta\mu/\mu \neq 0$, we would expect to see this problem more frequently, and more obviously, for transitions with larger $|K_i|$. The number of transitions rejected was small, and did not appear to be correlated with $|K_i|$, and hence it is unlikely that we are biasing $\Delta\mu/\mu$ towards zero.

3.2.3 Continuum and zero level fitting

Determination of the local quasar continuum is difficult in the Lyman- α forest due to absorption which can cover many Angstroms. Malec et al. (2010) discussed the fact that different investigators may manually fit qualitatively similar but quantitatively different continua to particular regions of a spectrum. For this reason, in any region where it appears that the local continuum has not been well determined, we allow for a freely floating continuum that is described by either a constant offset or a linear function, depending on the region in question. The use of a continuum fit which is at most linear helps to prevent strong degeneracies with the parameters of the Voigt profiles which might occur if a higher degree polynomial was used. The parameters of the continuum fit are determined simultaneously with all other parameters in the fit, and therefore any uncertainty in determining the continuum level naturally flows into the uncertainty on $\Delta\mu/\mu$.

Although weak night sky emission is subtracted as part of the initial flux extraction, the subtraction process does not appear to be optimal in the UVES CPL pipeline. In particular, in the base of saturated lines there appears to be a residual flux of about 2 percent of the local continuum. Therefore, in any region which includes saturated lines we allow for the zero level to be a free parameter in the fit, to be determined simultaneously with all other parameters.

3.2.4 Over-fitting

It is possible to add too many H I components to a particular region, leading to “over-fitting”. Over-fitting is undesirable for several reasons. The primary reason is that it means that another, simpler model can explain the absorption spectrum better than the over-

fitted model. Perhaps more importantly, it means that the performance of the optimisation algorithm can be substantially impaired. With significant over-fitting, convergence to the χ^2 minimum can be excessively slow. In extreme cases, convergence may not occur at all. Over-fitting can be detected through two means:

i) The addition of components which increase the AICC suggests that the components are not supported by the data. If the AICC significantly decreases upon removal of the components, this suggests that the model was over-fitted.

ii) Over-fitting causes the uncertainty estimates on the parameters of the components in question to be excessively large (Gill et al. 1986). In fact, this is often a good way to directly identify components which are potentially unnecessary; the AICC relates to the model as a whole and therefore cannot suggest which components may be unnecessary. In particular, H I transitions with $\sigma_{\log_{10} N} \gtrsim 1.0$ or $\sigma_b/b \gtrsim 1$ are certainly suspicious. In regions with substantial over-fitting, errors can easily be substantially larger than this. The numerical cause of these large errors is strong relative degeneracies between parameters. That is, χ^2 is almost flat in some direction in the parameter space relating to the offending transitions. It is this flatness in χ^2 which is the cause of poor convergence. Nevertheless, the presence of large errors on some components does not mean that they are unnecessary. In particular, the column densities for transitions which are saturated can be very poorly determined. This necessarily means that saturated H I transitions will have large errors on the column density.

Because of the impact of over-fitting on the convergence of VPFIT, we spent considerable effort trying to identify cases of over-fitting, and removed H I components as necessary to minimise the problem.

3.2.5 Determination of the number of velocity components in the Q0528–250 absorber

King et al. (2008) noted that at least 2 velocity components are plainly required to achieve an adequate fit to the Q0528–250 absorber. They argued on the basis of the AICC that 4 velocity components were required. We followed a similar procedure here, by comparing models with two, three and four velocity components. We choose whichever model has the lowest AICC (see section 4.2 below).

3.3 Voigt profile fits

Our final fits were obtained where we were not able to obtain any statistically appreciable improvement. We show one part of the H₂/HD and Lyman- α Voigt profile fit in Fig. 1; the full model may be found in the online-only version in Appendix A.

3.4 Determining $\Delta\mu/\mu$

There are two methods in the literature used for determining $\Delta\mu/\mu$ for a particular H₂ absorber. These can be described as the ‘line-by-line fitting method’ (LBLFM) and the ‘comprehensive fitting method’ (CFM).

In the LBLFM, for each transition one calculates the ζ_i values in equation 3. Equation 3 then implies that a plot of ζ_i vs K_i will have a slope of $\Delta\mu/\mu$. The slope of a linear fit to ζ_i vs K_i can easily be obtained from standard χ^2 minimisation techniques. This method has the advantage of assisting visual detection of problematic transitions, where ζ_i deviates from the general trend shown by

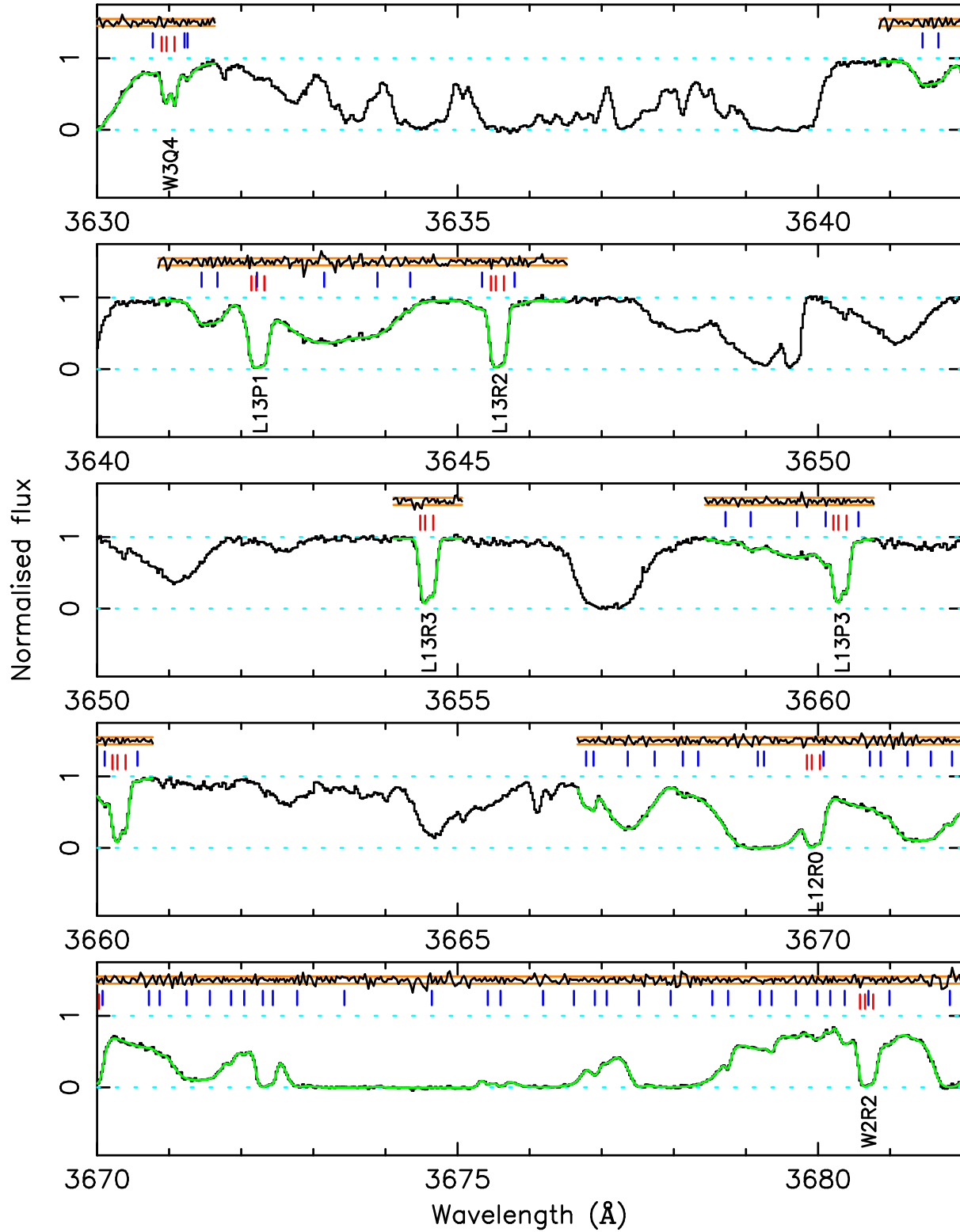


Figure 1. Part of the H₂/HD Voigt profile fit for the $z = 2.811$ absorber toward Q0528–250. The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data. Other regions shown in Appendix A (online only).

other lines. A strong deviance of this nature implies that a problem exists in the fit to the transition which generated that ζ_i value; one can then return to the spectral fit to attempt to determine the problem. This method was used by Ivanchik et al. (2005) and Reinhold et al. (2006).

Unfortunately, the LBLFM cannot be used when the absorber contains two or more closely spaced velocity components. The If velocity components overlap then the calculated errors in the line centroids for those velocity components will be correlated. It follows that, to use the LBLFM, one must first demonstrate that a single component is the best model for the H_2/HD lines i.e. that a two component model is statistically worse than a one component model. The LBLFM can be generalised to fit closely spaced velocity components through the use of generalised least squares methods, which allow for correlations between data points. However, to the best of our knowledge this has not been used in the literature to determine $\Delta\mu/\mu$ from a quasar spectrum.

The CFM allows one to reliably measure $\Delta\mu/\mu$ in the case where the absorber contains two or more closely spaced velocity components. In the CFM, one fits all H_2 transitions simultaneously with a single redshift for each velocity component, and allows for $\Delta\mu/\mu$ as a free parameter in the fit, such that the rest wavelengths are perturbed as $\lambda_i^0 \rightarrow \lambda_i^0[1 + K_i(\Delta\mu/\mu)]$. The value of $\Delta\mu/\mu$ at the χ^2 minimum is the best estimate of $\Delta\mu/\mu$ from the spectrum. In this way, the overlap of the velocity components is naturally accounted for within the fitting process, including the error estimate for $\Delta\mu/\mu$. Additionally, this method significantly reduces the number of free parameters in the fit, which should improve the reliability of the method. This method was used by King et al. (2008) and Malec et al. (2010).

We have assumed that the errors in the K_i values (which are ≈ 1 percent) are negligible. This assumption can be clearly justified by examining typical reduced redshift plots e.g. Fig. 1 in (King et al. 2008), where it is clear that the uncertainty in determining the velocity shifts of the line centroids is much more important than ≈ 1 percent errors in the K_i coefficients.

The absorber in Q0528–250 contains several closely spaced molecular velocity components. We therefore use the CFM to derive our estimate of $\Delta\mu/\mu$. We note that our fits to each region were all constructed assuming $\Delta\mu/\mu = 0$; $\Delta\mu/\mu$ was allowed to vary only once our model for the H_2/HD and Lyman- α lines was complete. If $\Delta\mu/\mu$ is significantly different from zero, it is conceivable that this method could bias our measurement of $\Delta\mu/\mu$ towards zero by “fitting away” any μ variation through the choice of the Lyman- α forest model. However, because the forest lines generally have high b parameters and because our model cannot fit the spectra in a totally arbitrary fashion (our model consists of a series of Voigt profiles, which have rather less flexibility to fit the spectrum than polynomials, for example) any bias induced should be small. Note that, on account of the above methodology, we cannot bias $\Delta\mu/\mu$ away from zero.

3.5 HD

HD is sensitive to a change in μ , and therefore here we include HD in our analysis of $\Delta\mu/\mu$. Although HD should display a similar velocity structure to H_2 , the low optical depth of the HD transitions and the small number of transitions observed means that any such structure is unresolved. We therefore model the HD absorption with only a single velocity component. That is, the constraint on $\Delta\mu/\mu$ from HD is derived only by considering potential velocity shifts of the HD lines with respect to each other, and not with respect

to H_2 transitions. The small number of HD transitions used means that the statistical constraint on $\Delta\mu/\mu$ derived from HD is weak compared to that derived from the H_2 transitions. However, future detected HD absorbers observed with greater optical depths and higher SNRs may provide statistically competitive constraints.

3.6 Spectral extraction problems

The CPL pipeline appears to incorrectly estimate the errors associated with the flux data points in the base of saturated lines. In particular, the dispersion of the flux data points about the local zero level is too large to be accounted for by the statistical error. The statistical errors on the flux points appear to be under-estimated by a factor of at least 2. Although it is difficult to determine precisely what happens in regions of low, but non-zero flux, we believe that the errors there are also underestimated. The effect of this is to give falsely high precision on any quantity derived from these data points (including $\Delta\mu/\mu$). Additionally, one cannot fit plausible models to data involving regions of low or negligible flux.

When we combine the individual exposures into a final spectrum using UVES_POPLER, the program calculates a check on the consistency of the exposures contributing to each pixel. As the combined value of each flux pixel is given as a weighted mean of pixels from the contributing spectra, this consistency check is a value of χ_ν^2 for each pixel about the weighted mean. We attempt to correct for the under-estimation of uncertainties in regions of low or negligible flux by applying the following algorithm. For each pixel in the combined spectrum, we take a region of five pixels centred on that point. We then take the median of the χ_ν^2 values just described that are associated with those five points. We then multiply the error estimate for that spectral pixel by the square root of that median value (that is, $\sigma_i \rightarrow \sigma_i \times \sqrt{\text{median}[\chi_\nu^2]}$).

However, we only increase the error array if $\text{median}(\chi_\nu^2) > 1$. This is because the spectral errors are generated from photon-counting, and therefore the statistical uncertainty should be a lower bound on the true uncertainty. Thus, it is difficult to justify decreasing the error array without particular evidence that the spectral uncertainty is systematically under-estimated.

4 RESULTS

In section 4.2 we describe our determination of the velocity structure of the absorber, and the $\Delta\mu/\mu$ values (and associated statistical errors) which result. In section 4.3 we describe various consistency checks we make to ensure that our results are robust. In section 5, we detail our analysis of potential systematic errors. In section 6 we give our preferred result including the final estimate of systematic errors.

4.1 Transitions used

For our analysis, we have included 76 molecular hydrogen transitions. In Fig. 2 we give the relationship of K_i and J with λ_0 for the H_2 and HD transitions used in our analysis. The seven HD transitions we have used are: L1R0, L2R0, L3R0, L4R0, L5R0, L6R0, L8R0. We show the spectral regions for the stronger HD transitions as a velocity plot in Fig. 3, clearly demonstrating the existence of HD.

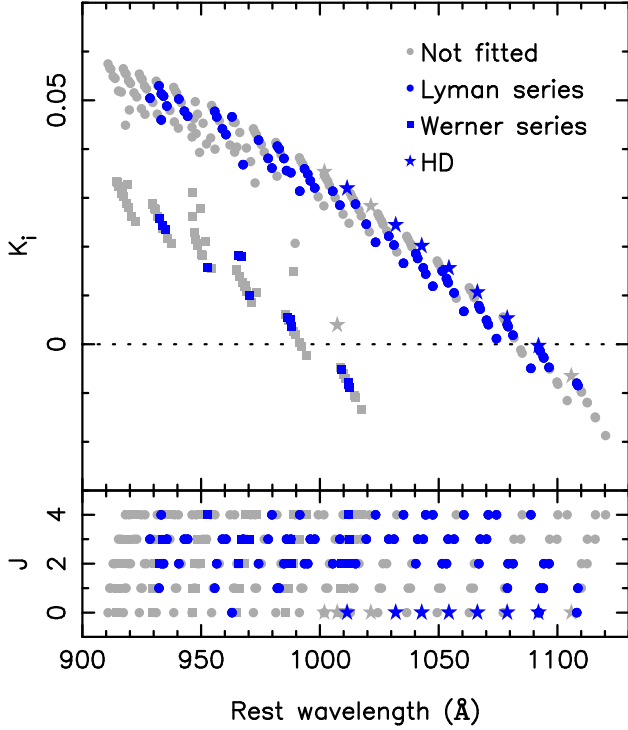


Figure 2. Relationship of K_i and J with λ_0 (rest wavelength) for the H_2 and HD transitions used in our analysis of the $z = 2.811$ absorber toward Q0528–250. Only H_2 transitions with $J \in [0, 4]$ are shown. All the HD transitions have $J = 0$. *Upper panel:* the sensitivity coefficients, K_i , for the transitions used in our analysis of the absorber (dark blue points) and not detected or not fitted (grey points). *Lower panel:* the distribution of transitions with wavelength according to their J -level.

4.2 Velocity structure and $\Delta\mu/\mu$ results

We examine the question of the velocity structure of the Q0528–280 absorber under different scenarios. To do this, we consider models with two, three and four velocity components. In the 3- and 4-component case, we applied a model in which corresponding velocity components in transitions from the same J -level had the same b parameter [$b = F(J)$], and also the scenario in which corresponding transitions had the same b parameter regardless of J [$b \neq F(J)$]. For the 2-component case, we only considered $b = F(J)$. We give the AICC for these scenarios, and the resulting values of $\Delta\mu/\mu$, in Table 2. We note that we were unable to obtain a stable fit for a 4-component model where b parameters for corresponding components were forced to be the same for all J -levels. In this model, the column density of one of the components in the $J = 1$ transitions was driven down to effectively zero. This component was the second strongest component in the $J = 2$ and $J = 3$ transitions. Although we could have omitted this component, the substantial differences in relative strength between the different components in the different J -levels means that this model is very unlikely to be a good representation of the physical situation, and therefore that the value of $\Delta\mu/\mu$ derived is unlikely to be accurate.

We note the sign change in $\Delta\mu/\mu$ from the result in King et al. (2008) [$\Delta\mu/\mu = (-1.4 \pm 3.9) \times 10^{-6}$], although in the $n = 3$ case the result is only marginally different from zero. It is clear from the $n = 3$ results that a model where different J -levels have different b parameters is preferred very strongly over a model with the same

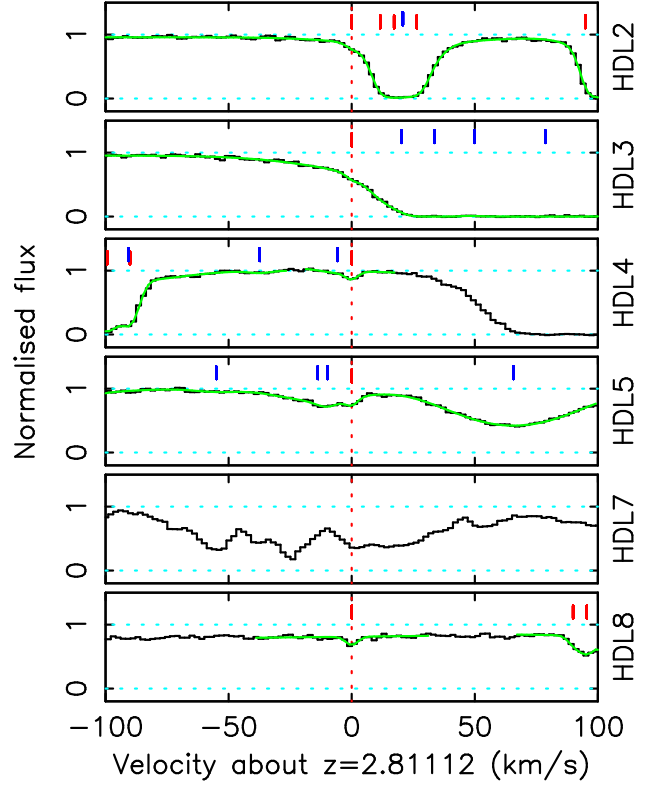


Figure 3. Different spectral regions for our spectrum of Q0528–250 demonstrating the presence of HD in the absorber at $z = 2.811$. The velocity is shown about $z = 2.81112$, which is the fitted HD redshift. The black line is the spectral data and the green line is the model fitted to the data. The red tick marks at $v = 0$ are fitted HD transitions. The red tick marks at $v \neq 0$ are fitted H_2 components. The blue tick marks are fitted Lyman- α components.

Table 2. Analysis of the velocity structure of Q0528–250. n gives the number of components. The second column defines whether the b -parameter for different components is fixed or is different for transitions with different J , i.e. whether b is a function of J [$b = F(J)$]. The column labelled AICC gives the AICC for the fit. The column Δ AICC shows the difference of the AICC with respect to the best-fitting model. For the 4-component, $b \neq F(J)$ model, the column density of one component of the $J = 1$ transitions was driven to zero, and the component was rejected, where this component was strongly detected in other J -levels. This implies that the model is not physically realistic, and so we label it as unstable. For these fits, the instrumental resolution was assumed to have a FWHM of 5.45 km s^{-1} . Note that the AICC is numerically quite large because of the thousands of degrees of freedom due to thousands of spectral points being included in the fit.

n	$b \stackrel{?}{=} F(J)$	AICC	Δ AICC	χ^2_ν	$\Delta\mu/\mu$ (10^{-6})
2	$b = F(J)$	11545.8	57.6	1.126	2.6 ± 2.8
3	$b = F(J)$	11488.2	0	1.115	0.2 ± 3.2
3	$b \neq F(J)$	11653.8	165.6	1.141	0.4 ± 3.2
4	$b = F(J)$	11510.4	22.2	1.117	3.4 ± 3.7
4	$b \neq F(J)$	Unstable fit	n/a	n/a	n/a

b -parameter for each J -level. This accords well with the results of King et al. (2008).

We note that, for these fits, the 3-component model is preferred to the 4-component model, at about the same statistical significance as the 4-component model was preferred to the 3-component model in King et al. (2008). There are several points to note here:

i) The seeing conditions for the spectra used by King et al. (2008) were quite variable, and may have induced a significantly non-Gaussian instrumental profile. The requirement for 4 components in King et al. (2008) may be a reflection of the non-Gaussian profile, rather than the absorber itself. Because the components are unresolved at these resolving powers, identification of the correct number of components is difficult. Nevertheless, what is important is that all the statistical structure in the measured profiles is adequately accounted for. This is directly reflected in the AICC.

ii) Some of the exposures contributing to the spectrum used by King et al. (2008) were poorly calibrated. It is conceivable that wavelength miscalibrations could cause a model with more complexity to be favoured.

iii) Although we have attempted to apply the same forest model in analysing the 3- and 4-component model (where in each case χ^2 is obviously minimised with respect to all the parameters), the construction of the forest model itself depends on the choice of the H_2 model. Strong H_2 components will obviously have little effect on the forest model because they are clearly distinguished from the forest. However, the 4th component of the model is weak and unresolved visually. We created the 3-component fit by removing the weakest component from the 4-component fit. In most regions, the resulting fit was reasonable, however in a small number of regions we found that we had to add weak forest components to account for the removal of the 4th H_2 component. To achieve a like-with-like comparison, we included these extra forest components in the 4-component fit. This means that any test for the statistical significance of the number of components depends somewhat on the choice of forest model near the H_2 lines.

iv) We noted whilst we were iteratively refining the model in the 3- and 4-component cases that the 4-component model was preferred for most of the refining process, with $\Delta\text{AICC} \approx 10$ in its favour for much of the time. It was only in the last few rounds of refining the model that the 3-component model became preferred as a result of changes made to a small number of regions. Therefore, the choice of the correct number of components can be sensitive to decisions made about the forest model in a small number of regions.

v) If the instrumental resolution used differs significantly from the true instrumental resolution then one might choose an incorrect number of components.

We conclude from this that although the Jeffreys' scale suggests that there is very significant evidence for the 3-component model over the 4-component model, in light of the fact that this evidence is conditioned on the correct choice of forest model and instrumental resolution, the actual preference for the 3-component model over the 4-component is rather weak. These arguments apply similarly to the preference for a 4-component model over a 3-component model in King et al. (2008). The issue of whether there are 3 or 4 components is simply very difficult to resolve given the actual SNR and resolution of the spectra available.

To investigate the effect of varying the instrumental resolution, we re-ran the 3- and 4-component, $b = F(J)$ fits with an instrumental resolution of 5.15 km s^{-1} . The results of this are given in Table 3. There are three points to note from varying the instru-

Table 3. Effect of varying the instrumental resolution from $\text{FWHM}=5.45 \text{ km s}^{-1}$ (table 2) to $\text{FWHM}=5.15 \text{ km s}^{-1}$ on estimates of $\Delta\mu/\mu$ for different models. n gives the number of components. The second column defines whether the b -parameter for different components is fixed or is different for transitions with different J , i.e. whether b is a function of J [$b = F(J)$]. The column labelled AICC gives the AICC for the fit. The column ΔAICC shows the difference of the AICC with respect to the best-fitting model.

n	$b \stackrel{?}{=} F(J)$	AICC	ΔAICC	$\Delta\mu/\mu (10^{-6})$
3	$b = F(J)$	11470.4	0	0.3 ± 3.2
4	$b = F(J)$	11500.1	29.8	4.6 ± 3.8

mental resolution: *i)* in both cases the AICC decreases, suggesting that the lower instrumental resolution is preferred; *ii)* the preference for the 3-component model over the 4-component model increases, and; *iii)* $\Delta\mu/\mu$ is relatively insensitive to the instrumental resolution, as one would naively expect (particularly for the 3-component model). We give the resulting parameters from this fit in Table 4.

In principle, one can account for the uncertainty in the choice of model by constructing a weighted mean of the $\Delta\mu/\mu$ values from each model, where the weights are given by the penalised likelihood [i.e. $w_i = \exp(-\text{AICC}/2)$]. Because the difference in the AICC between the 3-component and 4-component models is ≈ 30 , however, the suppression of the 4-component model is sufficiently large so that any contribution from this model can be neglected.

Thus, on the basis of the statistical results in Table 3, we choose $\Delta\mu/\mu = (0.3 \pm 3.2) \times 10^{-6}$ as our preferred statistical result for the analysis of $z = 2.811$ absorber toward Q0528–250. This is obviously consistent with no change in μ .

4.3 Consistency checks

We can relax some of the assumptions made in our analysis of this absorber to explore whether they have a meaningful impact on the result. In particular, we explore here whether the result we obtain is significantly affected by our assumptions about the different J -levels. We follow a similar procedure to that used by Malec et al. (2010). We also verify that our optimisation algorithm is functioning adequately.

4.3.1 Different $\Delta\mu/\mu$ from different J -levels

Rather than allowing transitions from all J -levels to contribute to a single value of $\Delta\mu/\mu$, within VPFIT we can calculate a value of $\Delta\mu/\mu$ for each H_2 J -level (and one for HD) separately. Strictly, the values of $\Delta\mu/\mu$ obtained are not independent because: *i)* they assume that each J -level has the same number of velocity components; and, *ii)* the redshifts are tied between corresponding velocity components in transitions arising from different J -levels. Nevertheless, this is useful for quantifying the contribution that each J -level makes to the final result. Ubachs et al. (2007) noted that, on account of the para-ortho distribution of H_2 the $J = 1$ state is significantly populated even at low temperatures. They suggested dividing the states into a $J \in [0, 1]$ set (cold states) and $J \geq 2$ set (warm states) to examine the impact of temperature. We examine both of these cases in Fig. 4. We see that there is no clear evidence for a difference of $\Delta\mu/\mu$ obtained using transitions arising from different J -levels.

Table 4. Best-fitting H₂/HD model parameters in the 3 velocity component model by J -level. n gives the number of H₂ transitions fitted for that J -level. The columns labeled “Component 1” through “Component 3” give the parameters for the 3 H₂ components, whilst the HD column gives the parameters for the single fitted HD component. For the H₂ components, N_{rel} specifies the relative strengths of the three components, defined as $N_{\text{rel},i} = N_i / \sum_j N_j$ for the components in that J -level; because we fit the line intensities as free parameters for the H₂ transitions it is difficult to derive an accurate estimate of the column density. For an approximate estimate of the total column density for each J -level, see Table 6. The HD components were fitted using the actual oscillator strengths, and so we give a direct estimate of the column density for the $J = 0$ level of HD explicitly. b gives the b parameter for the component. Uncertainty estimates are derived from the covariance matrix at the purported optimisation solution. For certain poorly-determined b parameters, where $\sigma_b/b \gtrsim 1$ these uncertainty estimates will not yield accurate confidence intervals on the b -parameters due to the fact that $b > 0$. Our optimiser has a hard lower limit for b of 0.4 km s^{-1} to prevent numerical problems with the convolution of the Voigt profile model with the instrumental profile model, which is why the b parameters for the $J = 3$ and $J = 4$ levels of component 3 are both 0.40 km s^{-1} . Note the trend for decreasing b with increasing J for the H₂ components, which is why the $b = F(J)$ models are significantly preferred over the $b \neq F(J)$ models.

J -level	n	Parameter	Component 1	Component 2	Component 3	HD
		z	2.8110056(9)	2.8111223(7)	2.8109346(11)	2.8111200(33)
$J = 0$	3	$b \text{ (km s}^{-1}\text{)}$	5.36 ± 0.36	3.56 ± 0.22	1.09 ± 0.94	0.95 ± 0.87
		N_{rel}	0.68	0.25	0.08	$\log_{10}(N/\text{cm}^{-2}) = 13.27 \pm 0.07$
		$b \text{ (km s}^{-1}\text{)}$	4.64 ± 0.99	3.57 ± 0.25	2.44 ± 0.62	
$J = 1$	8	N_{rel}	0.69	0.22	0.09	
		$b \text{ (km s}^{-1}\text{)}$	4.63 ± 0.25	3.05 ± 0.08	0.42 ± 0.08	
		N_{rel}	0.18	0.07	0.75	
$J = 2$	24	$b \text{ (km s}^{-1}\text{)}$	3.78 ± 0.19	3.08 ± 0.08	0.40 ± 0.05	
		N_{rel}	0.14	0.07	0.78	
		$b \text{ (km s}^{-1}\text{)}$	3.58 ± 0.26	1.88 ± 0.15	0.40 ± 0.22	
$J = 3$	28	N_{rel}	0.42	0.46	0.12	
		$b \text{ (km s}^{-1}\text{)}$				
		N_{rel}				

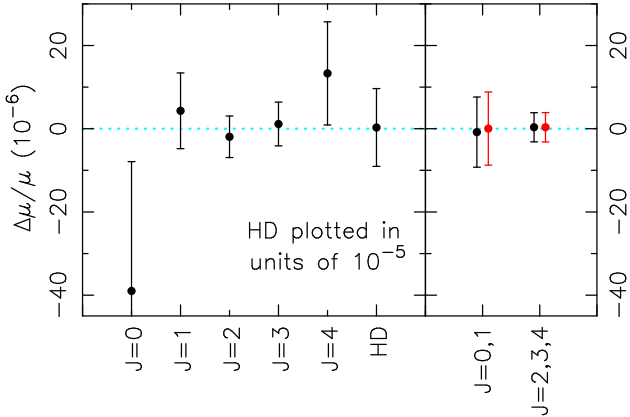


Figure 4. Left panel: $\Delta\mu/\mu$ for each H₂ J -level and HD, assuming that each J -level has the same 3-component velocity structure. The $J = 0$ level has a substantially larger uncertainty than the $J \in [1, 4]$ transitions because only 3 $J = 0$ transitions are used in the fit. HD has been plotted in units of 10^{-5} to increase clarity for the H₂ results. Right panel: $\Delta\mu/\mu$ for two groups of transitions, H₂ $J \in [0, 1]$ (“cold transitions”) and $J \in [2, 4]$ (“warm transitions”). The black data points are the results where the redshifts for corresponding components in the cold and warm transitions are assumed to be the same, and the red points are where the redshifts are allowed to differ between the cold and warm transitions.

4.3.2 Optimiser performance

The determination of $\Delta\mu/\mu$ requires minimising χ^2 for a model which contains hundreds of free parameters. Adequate performance of the optimiser is crucial in obtaining a reliable value of $\Delta\mu/\mu$ and an associated error. Premature termination of the optimiser will yield an erroneous best estimate for $\Delta\mu/\mu$. Typically this result will be biased towards the value of $\Delta\mu/\mu$ that is used to initialise the optimiser. Code errors or other gross problems with the optimiser will also render the error associated with $\Delta\mu/\mu$ unreliable.

In the context of potential variation of α , we have checked that the optimiser within VPFIT performs adequately for the case of a

few tens of parameters using Markov Chain Monte Carlo (MCMC) methods (King et al. 2009). We found for three systems that the estimated value of $\Delta\alpha/\alpha$ and its associated 1σ error provided by VPFIT are very similar to that determined from MCMC exploration of the likelihood function. Nevertheless, our model for the absorber in Q0528–250 and the Lyman- α forest have more than an order of magnitude more free parameters, and so an explicit check on the performance of the optimiser is advantageous.

A simple check on whether the optimiser is functioning adequately is to restart VPFIT with an initial value of $\Delta\mu/\mu$ that is significantly different to both the original starting guess ($\Delta\mu/\mu = 0$) and the final result returned by VPFIT. To do this, we re-ran the optimisation with a starting value of $\Delta\mu/\mu = 10^{-5}$ for the 3-component, $b = F(J)$ fit with an instrumental resolution FWHM of 5.45 km s^{-1} . The final value of $\Delta\mu/\mu$ under this circumstance was 0.210×10^{-6} , compared to 0.187×10^{-6} when started from $\Delta\mu/\mu = 0$. These two numbers differ by $\approx 0.007\sigma$, which is entirely negligible. This demonstrates both that the final result is insensitive to reasonable choices of the starting guess for $\Delta\mu/\mu$ and that our optimiser is functioning adequately.

We note that Malec et al. (2010) performed Monte Carlo simulations on realisations of their J2123–0050 spectrum and found consistency between the error on $\Delta\mu/\mu$ and the distribution of $\Delta\mu/\mu$ values. Similar simulations confirm the reliability of VPFIT in the context of $\Delta\alpha/\alpha$ (Murphy 2002; Murphy et al. 2003; King et al. 2009).

5 SYSTEMATIC ERRORS

Malec et al. (2010) discussed a number of potential systematic errors which could affect the measurement of $\Delta\mu/\mu$ from the analysis of H₂/HD absorbers. We investigate the same potential systematic errors here.

It is worth noting initially that a clear potential source of uncertainty arises from the quality of the wavelength calibration. Because the K_i values are well-correlated with wavelength within the Lyman and Werner series (but much less so when considering both

the Lyman and Werner series simultaneously), any effect which systematically expands or compresses the wavelength scale will mimic a change in μ . Any effect which only locally perturbs the wavelength calibration will have a random effect on $\Delta\mu/\mu$, and, providing enough transitions are used, the impact of such an effect will average to zero. We discuss wavelength calibration in the next two sections.

5.1 Known wavelength calibration errors due to uncertainties in the ThAr calibration

The calibration of the ThAr wavelength scale is not perfect; each of the ThAr transitions displays a residual velocity offset about the best-fit polynomial solution. The RMS of the residuals is $\sim 70 \text{ m s}^{-1}$ in the blue arm and $\sim 55 \text{ m s}^{-1}$ in the red arm. However, these fluctuations are random, and therefore will average out if a large number of H_2 transitions are used. Only systematic deviations from the true wavelength solution should appreciably affect the best estimate of $\Delta\mu/\mu$. There are fewer good ThAr lines in the blue end of the spectrum than in the red end, and therefore larger deviations of the wavelength solution from the true solution are possible.

Following the analysis of Murphy et al. (2007), the systematic deviation in the blue end of the spectrum relative to the red end of the forest has an upper limit of $\sim 20 \text{ m s}^{-1}$. The maximum K_i value used in the fit is 0.053, whilst the minimum is -0.009 . This implies that the maximum possible systematic due to this effect is given by $\delta(\Delta\mu/\mu) = (\Delta v/c)/\Delta K_i$, which is 1.1×10^{-6} . In reality, the effect is likely to be smaller than this as positive deviations should tend to cancel somewhat with negative deviations. However, how to reduce the effect is unclear; it may not simply scale as $1/\sqrt{N}$ (for N transitions). Therefore, we retain this estimate as the maximum possible systematic effect due to this cause.

5.2 Intra-order distortions of unknown origin

The path that the quasar light takes through the telescope is similar but not identical to that from the ThAr calibration lamp – the ThAr light fills the slit nearly uniformly, whilst the quasar light does not, and the ThAr light does not pass through the telescope optics. Due to the different light paths, the wavelength scale of the quasar light may be different to that of the ThAr light; the differences between them may appear as an apparent distortion of the wavelength scale. Both long range and short-range distortions are possible.

Griest et al. (2010) identified a pattern of distortion within echelle orders in Keck/HIRES spectra, such that the wavelength scale at the centre of echelle orders is distorted with respect to that at the echelle order edges. The peak-to-peak velocity distortion is $\sim 500 \text{ m s}^{-1}$ at $\sim 5600 \text{ \AA}$. The distortion was identified by comparing the calibration of a spectrum obtained using a ThAr exposure to that obtained using an I_2 absorption cell. The iodine cell is placed in the quasar light path, and the characteristic absorption spectrum is imprinted on the quasar spectrum. The use of an iodine cell therefore obviates the concern about optical path differences when using a ThAr lamp. Unfortunately, an I_2 cell is not useful for calibration of general quasar observations, because the iodine transitions cover only a relatively narrow part of the optical range, and because of the loss of flux from the quasar as a result of the use of the cell. The observed distortion pattern appears to be dependent on wavelength, and the distortion may be larger at longer wavelengths. The precise origin of the distortions is unknown, and similarly it is unknown to what extent the distortions remain constant

in time, and how they depend on extrinsic factors (e.g. telescope orientation, temperature, pressure and accuracy of quasar centering in the spectrograph slit). Therefore, it is not possible at present to adequately remove these distortions of the wavelength scale from observations.

Whitmore et al. (2010) identified a similar effect in VLT/UVES spectra, with a peak-to-peak velocity distortion of $\sim 200 \text{ m s}^{-1}$. The distortion appears to be much less consistent between echelle orders than that seen by Griest et al., however. Further observations by one of us (MTM) have shown that the observed wavelength distortion is definitely not constant over long periods of time i.e. more than several nights, which makes removal of the distortion extremely difficult.

Similar to Malec et al. (2010), we attempted to estimate the magnitude of the error introduced into a determination of $\Delta\mu/\mu$ as a result of the observed velocity distortions. To do this, we used a triangular-shaped distortion, where the wavelengths of pixels at the centre of echelle orders were displaced by $+200 \text{ m s}^{-1}$ with respect to those at the echelle order edges. The modification to the spectra was implemented within UVES_POPLER. The shift in $\Delta\mu/\mu$ after modifying the spectrum was -0.3×10^{-6} [using the 3-component, $b = F(J)$ model]. Clearly this value is model-dependent – if the distortion has a different amplitude or form, then the impact on $\Delta\mu/\mu$ may be different. However, this estimate of the systematic error is likely to be of the correct magnitude. We therefore adopt a Gaussian with $\sigma = 0.3 \times 10^{-6}$ as an estimate of the systematic effect due to distortions of this type.

5.3 Velocity structure and spatial segregation

It is possible that transitions arising from different J -levels might be spatially segregated (Jenkins & Peimbert 1997; Levshakov et al. 2002). Assuming that all J -levels arise from the same redshift in this event could spuriously produce $\Delta\mu/\mu \neq 0$. Although our result for the Q0528–250 absorber is statistically consistent with zero, it is of course possible that a non-zero $\Delta\mu/\mu$ could be pushed towards zero by this sort of systematic effect. Similar to Malec et al. (2010), we relaxed our assumption that corresponding components in all J -levels arise from the same redshift. In particular, we divided the data set into “cold” transitions, $J \in [0, 1]$, and “warm” transitions, $J \in [2, 4]$, as was done earlier, but only tie the redshifts of corresponding components between different J -levels within these two groups. If there is spatial segregation, this should be seen as a statistically significant difference between the redshifts of corresponding components between the two groups, and also as a substantial shift in the values of $\Delta\mu/\mu$ derived from the two groups compared to what was obtained earlier.

In the right panel of Fig. 4 we directly compared $\Delta\mu/\mu$ in the case where the velocity structure was allowed to vary between cold and warm components, and note that there is no appreciable shift. In Fig. 5 we show these considerations more directly by examining the differences in the redshifts of the three components, and also the explicit difference between $\Delta\mu/\mu$ in the two cases considered. We see that there is no statistically significant difference between the redshifts in any of the three components when the velocity structure is allowed to vary between the cold and warm components. Similarly, we see that the shift in $\Delta\mu/\mu$ is $\lesssim 0.1\sigma$ when the velocity structure is allowed to vary between the cold and warm components. Thus, we conclude that there is no evidence for a systematic shift in $\Delta\mu/\mu$ as a result of segregation of the cold and warm J -levels.

Nevertheless, to quantify the possible error introduced by our

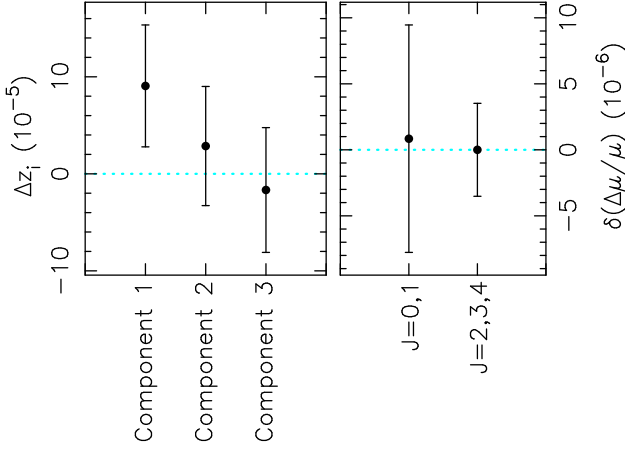


Figure 5. *Left panel:* Difference between the redshifts of components for the H_2 fit for the 3-component model to the absorber when the velocity structure was allowed to vary between the “cold” and “warm” components. The difference for each component is defined as $z(J \in [0, 1]) - z(J \in [2, 4])$. *Right panel:* The difference in $\Delta\mu/\mu$ when the velocity structure was allowed to vary, defined as $\Delta\mu/\mu(\text{structure allowed to vary}) - \Delta\mu/\mu(\text{structure not allowed to vary})$. The error estimate is calculated as the mean of the error estimates in the two cases considered; the errors do not differ appreciably between the two cases. From consideration of both of these plots there does not appear to be any significant evidence that our assumptions about the velocity structure model have affected our estimate of $\Delta\mu/\mu$.

assumptions regarding velocity structure, we examine the actual shifts in $\Delta\mu/\mu$. The shift in $\Delta\mu/\mu$ for the $J \in [0, 1]$ levels from the original value [for the 3-component, $b = F(J)$ model] is 0.8×10^{-6} , and for the $J \in [2, 4]$ levels is 0.2×10^{-6} . We thus take 0.8×10^{-6} as an estimate of the potential error introduced into our analysis due to assumptions about the velocity structure in order to be conservative.

5.4 Re-dispersion of spectra

The spectrum used is the result of the co-addition of exposures taken at different times, each with its own ThAr calibration. During the co-addition, the spectra are placed on a common wavelength grid. Because the spectra are re-binned, the choice of wavelength grid introduces correlations between neighbouring pixels. More importantly, the choice of the wavelength grid has the potential to affect $\Delta\mu/\mu$ by slightly distorting the absorption profile shapes. To investigate this, we examined the effect of shifting the wavelength grid by -0.2 , -0.1 , 0.1 and 0.2 pixels. The shifts this induced in $\Delta\mu/\mu$ from the original value are -1.3×10^{-6} , -2.1×10^{-6} , $+1.0 \times 10^{-6}$ and $+0.2 \times 10^{-6}$. The standard deviation of these values is $\approx 1.4 \times 10^{-6}$, and so we adopt 1.4×10^{-6} as an estimate of the potential error in $\Delta\mu/\mu$ on account of the re-dispersion of the contributing exposures.

6 RESULT INCLUDING SYSTEMATIC ERRORS

In Table 5, we accumulate the potential systematic errors from our discussion above and give our final estimate of $\Delta\mu/\mu$ including the systematic component. Although the distribution of systematic errors is likely to be Gaussian in many cases, the impact on $\Delta\mu/\mu$ arising from the distortion in the wavelength scale (for instance) is an upper limit. The probability distribution of the sum of random

variables is given by the convolution of their individual probability density functions. Thus, to estimate our final uncertainty, we convolve the distributions assumed for each of the sources of uncertainty, and give the standard deviation of the resultant distribution as our uncertainty estimate.

This yields our final estimate of $\Delta\mu/\mu$ for the $z = 2.811$ absorber toward Q0528–250 as

$$\frac{\Delta\mu}{\mu} = (0.3 \pm 3.2_{\text{stat}} \pm 1.9_{\text{sys}}) \times 10^{-6} \quad (5)$$

$$= (0.3 \pm 3.7)_{\text{tot}} \times 10^{-6}. \quad (6)$$

7 OTHER ANALYSIS

7.1 H_2/HD column density ratio

The HD/H_2 column density ratio is of astrophysical interest. We calculate the total H_2 column density under the 3-parameter, $b = F(J)$ model via the method given in the caption to Table 6 to be $\log_{10}(N/\text{cm}^{-2}) = 16.556 \pm 0.024$. Column densities for each of the H_2 J -levels are given in Table 6. The HD $J = 0$ column density is $\log_{10}(N/\text{cm}^{-2}) = 13.267 \pm 0.072$ directly from the fit. We assume that the total H_2/HD column density ratio is approximated by the total $\text{H}_2/\text{HD}(J = 0)$ column density ratio (Tumlinson et al. 2010). This directly gives the HD/H_2 column density ratio as $(5.4 \pm 1.1) \times 10^{-4}$, or $\log_{10}[N(\text{HD})/N(\text{H}_2)] = -3.3 \pm 0.2$. Alternatively, $\log_{10}[N(\text{HD})/2N(\text{H}_2)] = -3.6 \pm 0.2$. This compares with $\log_{10}[N(\text{HD})/2N(\text{H}_2)] = -4.1 \pm 0.2$ from the $z = 2.059$ absorber toward J2123–0050 and -4.8 ± 1.5 from the $z = 2.627$ absorber toward FJ0812+32B (Tumlinson et al. 2010). The result presented here for $\log_{10}[N(\text{HD})/2N(\text{H}_2)]$ from Q0528–250 differs from the value from J2123–0050 at the 1.8σ level, which is large but not grossly inconsistent, particularly since the estimate here is statistical only and does not consider gas cloud kinematics or potential systematic effects. Additionally, because the low- J H_2 transitions in the Q0528–250 absorber are saturated it is difficult to accurately estimate the column density for H_2 in Q0528–250; Tumlinson et al. (2010) noted this for the absorber in J2123–0050. As such, some caution is warranted in comparing the HD/H_2 ratio presented here to other results. Additionally, our fits were constructed with the H_2 line intensities fitted as free parameters. This makes it challenging to accurately estimate the total H_2 column density; to obtain a more accurate HD/H_2 column density ratio, this absorber should be re-analysed specifically for this purpose.

8 DISCUSSION

Firstly, it is clear that the result obtained here is consistent with no variation of μ . Secondly, the result is consistent with the Q0528–250 result in King et al. (2008), and also the $\Delta\mu/\mu$ estimates obtained from Q0405–443 and Q0347–383 in King et al. (2008) and from J2123–0050 in Malec et al. (2010). From consideration of all the H_2/HD measurements from quasar absorbers (which are all at $z > 2$) there is no statistically significant evidence for cosmological variation in μ . In particular, we can formally combine the results from King et al. (2008), Malec et al. (2010) and this work under a weighted mean model (assuming that a weighted mean model is a legitimate description of the data), and we obtain $\Delta\mu/\mu = (2.3 \pm 2.2) \times 10^{-6}$, with χ^2_ν about the weighted mean of ≈ 0.9 . The fact that $\chi^2_\nu \approx 1$ suggests that the uncertainty estimates on $\Delta\mu/\mu$ are approximately correct. We show the result of

Table 5. Error budget for the analysis of the $z = 2.811$ absorber toward Q0528–250 from the sources described in section 5. The third column gives the magnitude of the uncertainty estimate. The final error estimate is calculated as the standard deviation of the convolution of the assumed distributions of the individual error estimates; the assumed distributions are given in the fourth column.

Source of error	$\Delta\mu/\mu$ (10^{-6})	$\delta(\Delta\mu/\mu)$ (10^{-6})	Assumed distribution
Systematic distortion of ThAr wavelength scale		± 1.1	Uniform
Intra-order wavelength scale distortions		± 0.3	Gaussian
Velocity structure and spatial segregation		± 0.8	Gaussian
Re-dispersion of spectra		± 1.4	Gaussian
Total systematic error		± 1.9	
Statistical	0.3	± 3.2	Gaussian
Final estimate	0.3	± 3.7	

Table 6. Column densities for the different J -levels under the 3-component, $b = F(J)$ model. n gives the number of transitions contributing to a particular J -level. Because the column densities for each transition are fitted as free parameters, the \log_{10} column density for each J -level is calculated as the weighted mean of the \log_{10} column densities for the transitions in that J -level, but where statistical errors are increased in quadrature with an additional term, σ_J , which is calculated so that $\chi^2_\nu = 1$ about the weighted mean. Without the inclusion of σ_J , $\chi^2_\nu \gg 1$, which reflects the accumulation of errors in the local model for the continuum and weak, unmodelled Lyman- α transitions which overlap with the transitions in question. σ_J is given in the fourth column.

J	n	$\log_{10}(N/\text{cm}^{-2})$	$\log_{10}(\sigma_J)$
0	3	15.68 ± 0.25	0.43
1	8	16.028 ± 0.094	0.16
2	24	16.008 ± 0.055	0.23
3	28	15.917 ± 0.030	0.14
4	13	14.476 ± 0.025	0.084
Total	76	16.556 ± 0.024	n/a

this work in comparison with other quasar constraints on $\Delta\mu/\mu$ in Fig. 6.

9 CONCLUSION

We have used new observations of the molecular absorber at $z = 2.811$ towards Q0528–250 to derive a new, strong constraint on $\Delta\mu/\mu$ at high redshift, namely $\Delta\mu/\mu = (0.3 \pm 3.2_{\text{stat}} \pm 1.9_{\text{sys}}) \times 10^{-6} = (0.3 \pm 3.7) \times 10^{-6}$. This result is the strongest individual constraint on variation in μ at $z > 1$, and together with other measurements of H₂/HD absorbers provides a very strong constraint on evolution in μ over most of the observable history of the universe. The value of $\Delta\mu/\mu$ presented here is consistent with the values presented in King et al. (2008) and Malec et al. (2010).

Our result was derived using 76 H₂ transitions and seven HD transitions. This is the first time that HD has been detected in this absorber, with a column density of $\log_{10}(N/\text{cm}^{-2}) = 13.27 \pm 0.07$. We measure the H₂/HD column density ratio to be $\log_{10}[N(\text{HD})/N(\text{H}_2)] = -3.3 \pm 0.2$.

Similar to King et al. (2008) and Malec et al. (2010), we applied the comprehensive fitting method (CFM), where we simultaneously fitted the H₂/HD transitions with the surrounding Lyman- α forest lines. By modelling the forest simultaneously with the H₂/HD transitions, our estimate of $\Delta\mu/\mu$ should be more accurate. Similarly, the uncertainty in determining the locations of the forest lines directly feeds into our uncertainty estimate on $\Delta\mu/\mu$, helping to ensure that our uncertainty is not under-estimated.

On the basis of the AICC, we concluded that a model of the absorber with 3 velocity components is preferred. We imposed a

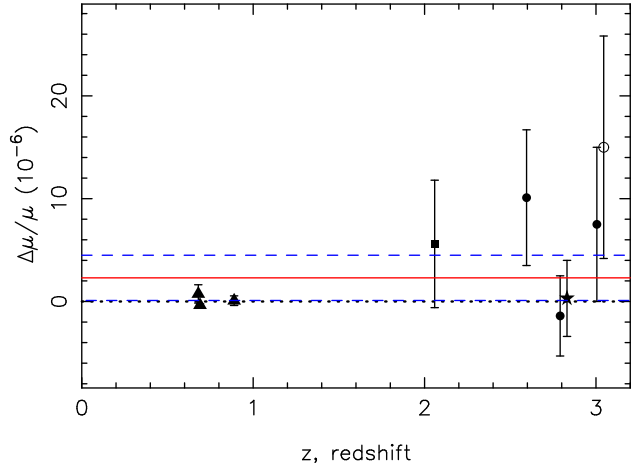


Figure 6. Recent cosmological constraints on $\Delta\mu/\mu$ derived from quasar absorption measurements. Filled circles indicate the H₂ measurements from Q0405–443, Q0347–383 and Q0528–250 in King et al. (2008), the square indicates the H₂/HD measurements from J2123–0050 in Malec et al. (2010), the star indicates the result of this paper, the open circle indicates the Q0347–383 measurement of Wendt & Molaro (2011) [which is not independent of the Q0347–383 measurement in King et al. (2008) due to the use of common spectra] and the triangles indicate the constraints from ammonia in Murphy et al. (2008a), Henkel et al. (2009) and Kanekar (2011). The two Q0528–250 points and the two Q0347–383 points have been slightly displaced in redshift for clarity. Where uncertainty estimates were given with both a statistical and systematic component, we added these in quadrature to derive the error bars shown. The red, solid line indicates the weighted mean of the $z > 1$ points excluding the Q0347–383 point from Wendt & Molaro (2011) [as it is not independent of the Q0347–383 point from King et al. (2008)], and the blue, dashed line indicates the 1σ uncertainty around that weighted mean. Clearly, if the $z < 1$ measurements were included in the weighted mean then they would dominate the result due to their small error bars.

number of assumptions about the velocity structure of the absorber in order to reduce the number of free parameters in the fit. Investigation of potential differences between ‘cold’ ($J \in [0, 1]$) and ‘warm’ ($J \in [2, 4]$) components showed that the potential impact of our assumptions on the estimate of $\Delta\mu/\mu$ is significantly smaller than the statistical error on $\Delta\mu/\mu$.

We investigated a number of potential systematic effects, including: known wavelength calibration errors due to uncertainties in the ThAr calibration; observed intra-echelle order wavelength scale distortions of unknown origin; assumptions about the velocity structure, and; the effect of re-dispersing the individual exposures onto a common wavelength grid. The contribution of each of these effects is small, and in aggregate the total uncertainty in $\Delta\mu/\mu$ is dominated by statistical rather than systematic error sources. This

is reassuring, and means that future VLT/UVES observations of H₂/HD absorbers may continue to increase the constraint on $\Delta\mu/\mu$ in a meaningful way, provided that appropriate absorbers can be identified.

Although there is no significant evidence for evolution in μ , and little ($\sim 2\sigma$) evidence for dipolar variation in μ across the sky, it is intriguing that the dipole direction in a dipole+monopole model fitted to the $z > 1$ H₂/HD $\Delta\mu/\mu$ constraints points in a similar direction to the $\Delta\alpha/\alpha$ dipole described in King et al. (2011). Given the distribution of existing H₂ results on the sky, new H₂ absorbers in judicious locations will be able to place strong constraints on spatial variation in μ .

The current number of quasars known to contain molecular hydrogen absorption is small, and almost all existing analyses have focused on the objects Q0405–443, Q0347–383, Q0528–250 and J2123–0050. Given that analysis of molecular hydrogen currently remains the most direct and precise way of investigating potential evolution in μ at $z > 2$, it would be advantageous if many more H₂/HD systems could be discovered. Although new observations of existing H₂/HD systems can increase the SNR, common systematics may remain. The use of more absorbers in different quasars would increase both the redshift and spatial coverage of the $\Delta\mu/\mu$ constraints.

ACKNOWLEDGMENTS

This work is based on observations carried out at the European Southern Observatory (ESO) under program ID 82.A-0087 (PI Ubachs), with the UVES spectrograph installed at the Kueyen UT2 on Cerro Paranal, Chile. The authors thank L. Kaper (Amsterdam), V. Flambaum (UNSW) and J. Berengut (UNSW) for fruitful discussions. JAK was supported in part by an Australian Postgraduate Award. JAK additionally thanks VU for support to travel to Amsterdam, where part of this work was carried out. MTM thanks the Australian Research Council for a QEII Research Fellowship (DP0877998). WU acknowledges support from the Netherlands Foundation for Fundamental Research of Matter (FOM).

This paper has been typeset from a \LaTeX file prepared by the author.

REFERENCES

- Abgrall H., Roueff E., 2006, *A&A*, 445, 361
 Abgrall H., Roueff E., Drira I., 2000, *A&AS*, 141, 297
 Abgrall H., Roueff E., Launay F., Roncin J., 1994, *Can. J. Phys.*, 72, 856
 Abgrall H., Roueff E., Launay F., Roncin J. Y., Subtil J. L., 1993, *A&AS*, 101, 273
 Akaike A., 1974, *IEEE Trans. Automat. Contr.*, 19, 716
 Bailly D., Salumbides E. J., Vervloet M., Ubachs W., 2009, *Mol. Phys.*, 108, 827
 Cowie L. L., Songaila A., 1995, *ApJ*, 453, 596
 D’Odorico S., Cristiani S., Dekker H., Hill V., Kaufer A., Kim T., Primas F., 2000, in J. Bergeron ed., *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series Vol. 4005 of Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Performance of UVES, the echelle spectrograph for the ESO VLT and highlights of the first observations of stars and quasars*. pp 121–130
 Flambaum V. V., Kozlov M. G., 2007, *Phys. Rev. Lett.*, 98, 240801
 Foltz C. B., Chaffee Jr. F. H., Black J. H., 1988, *ApJ*, 324, 267
 Gill P. E., Murray W., Wright M. H., 1986, *Practical optimization*. Elsevier, London
 Griest K., Whitmore J. B., Wolfe A. M., Prochaska J. X., Howk J. C., Marcy G. W., 2010, *ApJ*, 708, 158
 Henkel C., Menten K. M., Murphy M. T., Jethava N., Flambaum V. V., Braatz J. A., Muller S., Ott J., Mao R. Q., 2009, *A&A*, 500, 725
 Hollenstein U., Reinhold E., de Lange C. A., Ubachs W., 2006, *J. Phys. B.*, 39, L95
 Ivanchik A., Petitjean P., Varshalovich D., Aracil B., Srianand R., Chand H., Ledoux C., Boissé P., 2005, *A&A*, 440, 45
 Ivanov T. I., Dickenson G. D., Roudjane M., Oliveira N. D., Joyeux D., Nahon L., Tcham-Brillet W., Ubachs W., 2010, *Mol. Phys.*, 108, 771
 Ivanov T. I., Roudjane M., Vieitez M. O., de Lange C. A., Tcham-Brillet W., Ubachs W., 2008a, *Phys. Rev. Lett.*, 100, 093007
 Ivanov T. I., Vieitez M. O., de Lange C. A., Ubachs W., 2008b, *J. Phys. B: At. Mol. Phys.*, 41, 035702
 Jansen P., Xu L., Kleiner I., Ubachs W., Bethlem H. L., 2011, *Phys. Rev. Lett.*, 106, 100801
 Jeffreys H., 1961, *Theory of Probability*, 3 edn. Oxford University Press
 Jenkins E. B., Peimbert A., 1997, *ApJ*, 477, 265
 Kanekar N., 2011, *ApJ*, 728, L12
 King J. A., et al., 2011, *MNRAS*, submitted
 King J. A., Mortlock D. J., Webb J. K., Murphy M. T., 2009, *Mem. Soc. Astron. Italiana*, 80, 864
 King J. A., Webb J. K., Murphy M. T., Carswell R. F., 2008, *Phys. Rev. Lett.*, 101, 251304
 Kozlov M. G., Levshakov S. A., 2011, *ApJ*, 726, 65
 Levshakov S. A., Dessauges-Zavadsky M., D’Odorico S., Molaro P., 2002, *MNRAS*, 333, 373
 Liddle A. R., 2007, *MNRAS*, 377, L74
 Malec A. L., Buning R., Murphy M. T., Milutinovic N., Ellison S. L., Prochaska J. X., Kaper L., Tumlinson J., et al., 2010, *MNRAS*, 403, 1541
 Meshkov V. V., Stolyarov A. V., Ivanchik A. V., Varshalovich D. A., 2006, *Sov. J. Exp. Theor. Phys. Lett.*, 83, 303
 Murphy M., 2002, PhD thesis, University of New South Wales
 Murphy M. T., et al., 2001, *MNRAS*, 327, 1208
 Murphy M. T., Flambaum V. V., Muller S., Henkel C., 2008a, *Science*, 320, 1611
 Murphy M. T., Flambaum V. V., Webb J. K., Dzuba V., Prochaska J. X., Wolfe A. M., 2004, in S. G. Karshenboim & E. Peik ed., *Astrophysics, Clocks and Fundamental Constants Vol. 648 of Lecture Notes in Physics*, Berlin Springer Verlag, *Constraining Variations in the Fine-Structure Constant, Quark Masses and the Strong Interaction*. pp 131–150
 Murphy M. T., Tzanavaris P., Webb J. K., Lovis C., 2007, *MNRAS*, 378, 221
 Murphy M. T., Webb J. K., Flambaum V. V., 2003, *MNRAS*, 345, 609
 Murphy M. T., Webb J. K., Flambaum V. V., 2008b, *MNRAS*, 384, 1053
 Philip J., Sprengers J. P., Pielage T., de Lange C. A., Ubachs W., Reinhold E., 2004, *Can. J. Chem.*, 82, 713
 Potekhin A. Y., Ivanchik A. V., Varshalovich D. A., Lanzetta K. M., Baldwin J. A., Williger G. M., Carswell R. F., 1998, *ApJ*, 505, 523
 Reinhold E., Buning R., Hollenstein U., Ivanchik A., Petitjean P.,

- Ubachs W., 2006, *Phys. Rev. Lett.*, 96, 151101
- Salumbides E. J., Bailly D., Khrarov A., Wolf A. L., Eikema K. S. E., Vervloet M., Ubachs W., 2008, *Phys. Rev. Lett.*, 101, 223001
- Sugiura N., 1978, *Commun. Stat. A-Theor.*, 7, 13
- Thompson R. I., 1975, *Astron. Lett.*, 16, 3
- Thompson R. I., Bechtold J., Black J. H., Eisenstein D., Fan X., Kennicutt R. C., Martins C., Prochaska J. X., Shirley Y. L., 2009, *ApJ*, 703, 1648
- Tumlinson J., Malec A. L., Carswell R. F., Murphy M. T., Buning R., Milutinovic N., Ellison S. L., Prochaska J. X., Jorgenson R. A., Ubachs W., Wolfe A. M., 2010, *ApJ*, 718, L156
- Ubachs W., Buning R., Eikema K. S. E., Reinhold E., 2007, *J. Mol. Spectrosc.*, 241, 155
- Ubachs W., Reinhold E., 2004, *Phys. Rev. Lett.*, 92, 101302
- van Weerdenburg F., Murphy M. T., Malec A. L., Kaper L., Ubachs W., 2011, *Physical Review Letters*, 106, 180802
- Varshalovich D. A., Levshakov S. A., 1993, *J. Exp. Theor. Phys. Lett.*, 58, 237
- Varshalovich D. A., Potekhin A. Y., 1995, *Space Sci. Rev.*, 74, 259
- Webb J. K., et al., 2011, *Phys. Rev. Lett.*, submitted
- Webb J. K., Flambaum V. V., Churchill C. W., Drinkwater M. J., Barrow J. D., 1999, *Phys. Rev. Lett.*, 82, 884
- Webb J. K., Murphy M. T., Flambaum V. V., Dzuba V. A., Barrow J. D., Churchill C. W., Prochaska J. X., Wolfe A. M., 2001, *Phys. Rev. Lett.*, 87, 091301
- Wendt M., Molaro P., 2011, *A&A*, 526, A96
- Wendt M., Reimers D., 2008, *Eur. Phys. J. Special Topics*, 163, 197
- Whitmore J. B., Murphy M. T., Griest K., 2010, *ApJ*, 723, 89

APPENDIX A: VOIGT PROFILE FITS

Online only: Figures A1 through A14 show our Voigt profile model for the $z = 2.811$ absorber toward Q0528–250 and the surrounding Lyman- α forest regions, indicating both the positions of the H₂/HD components as well as the H I components used to fit the surrounding Lyman- α forest.

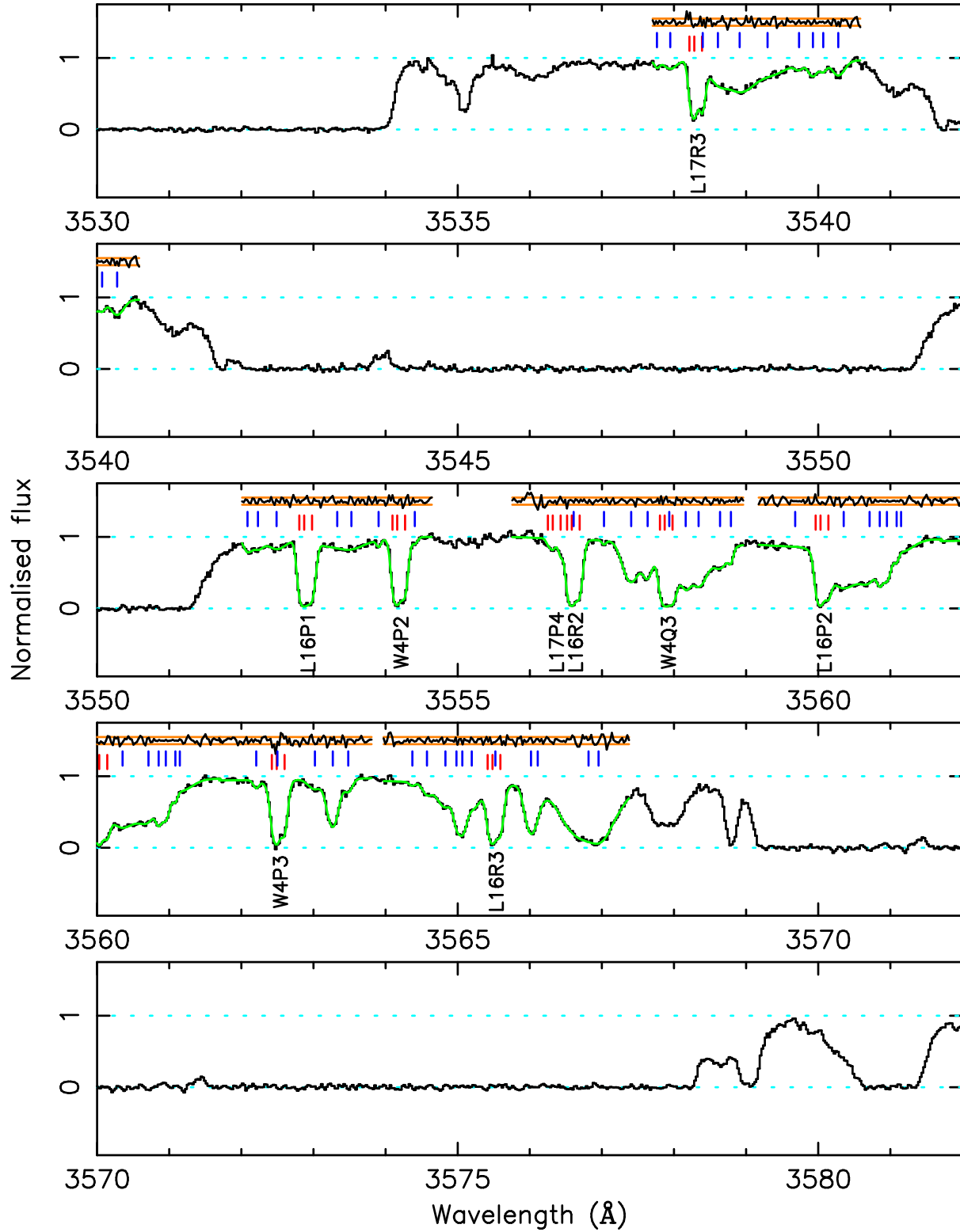


Figure A1. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 1). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

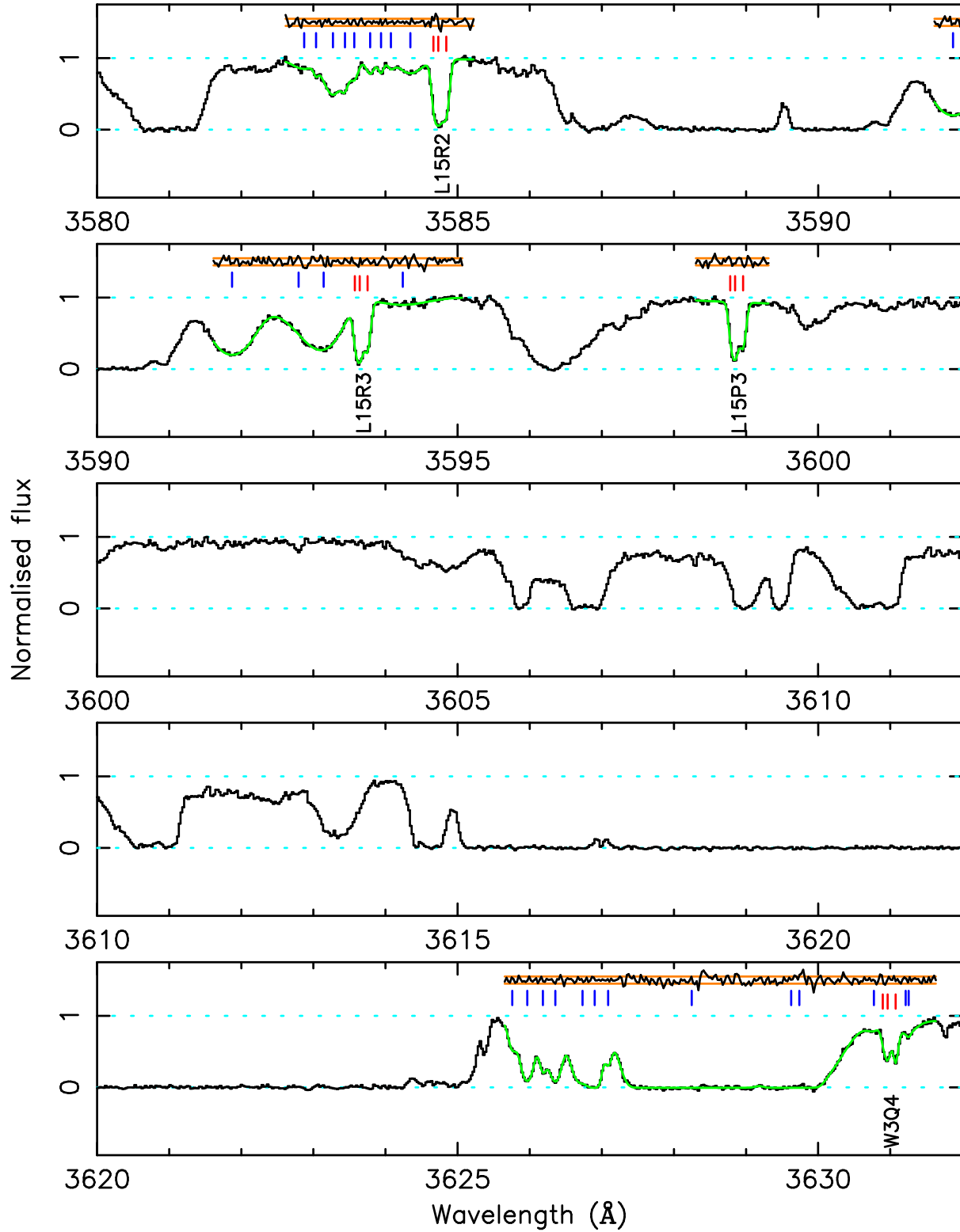


Figure A2. H_2/HD fit for the $z = 2.811$ absorber toward Q0528-250 (part 2). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

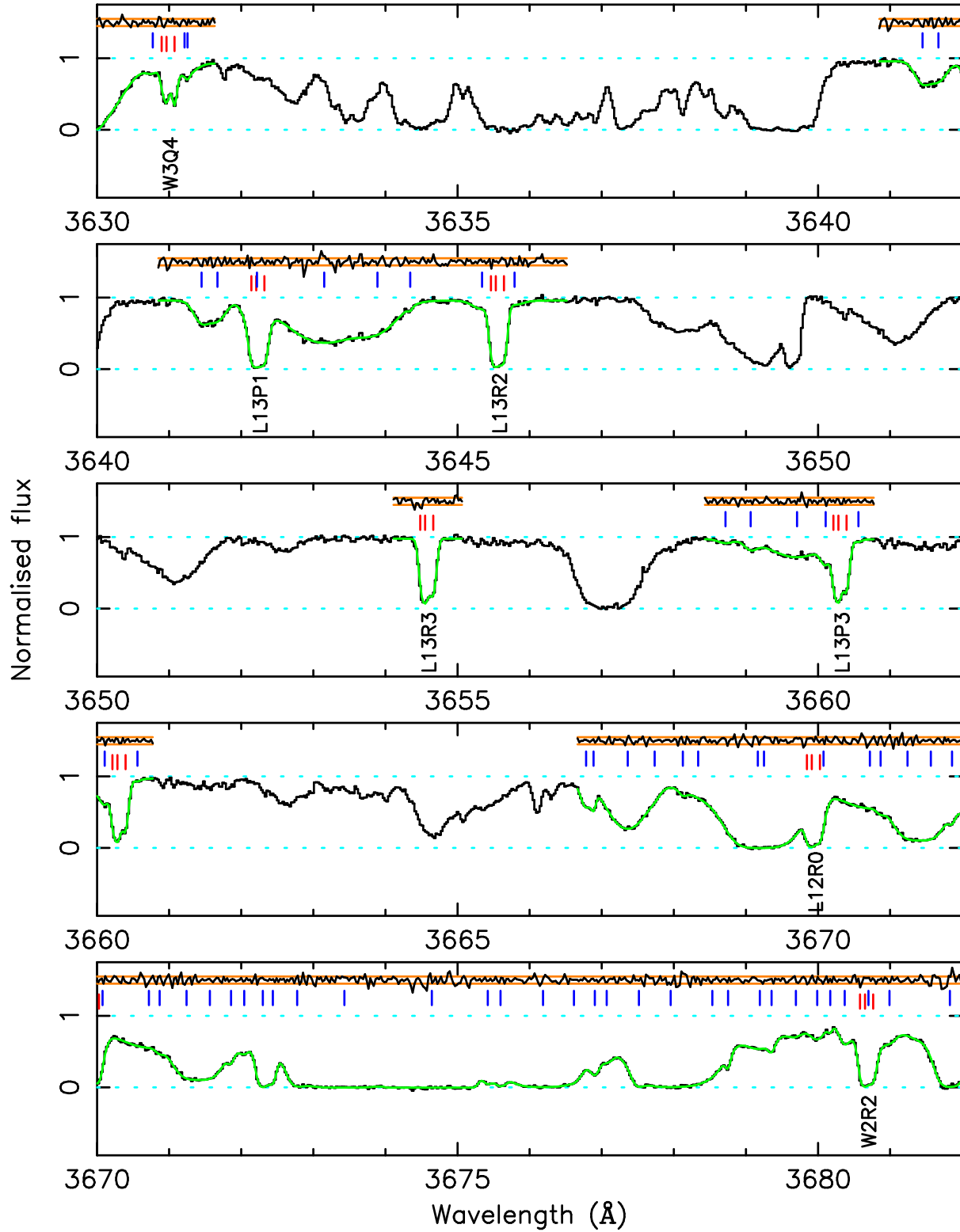


Figure A3. H₂/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 3). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

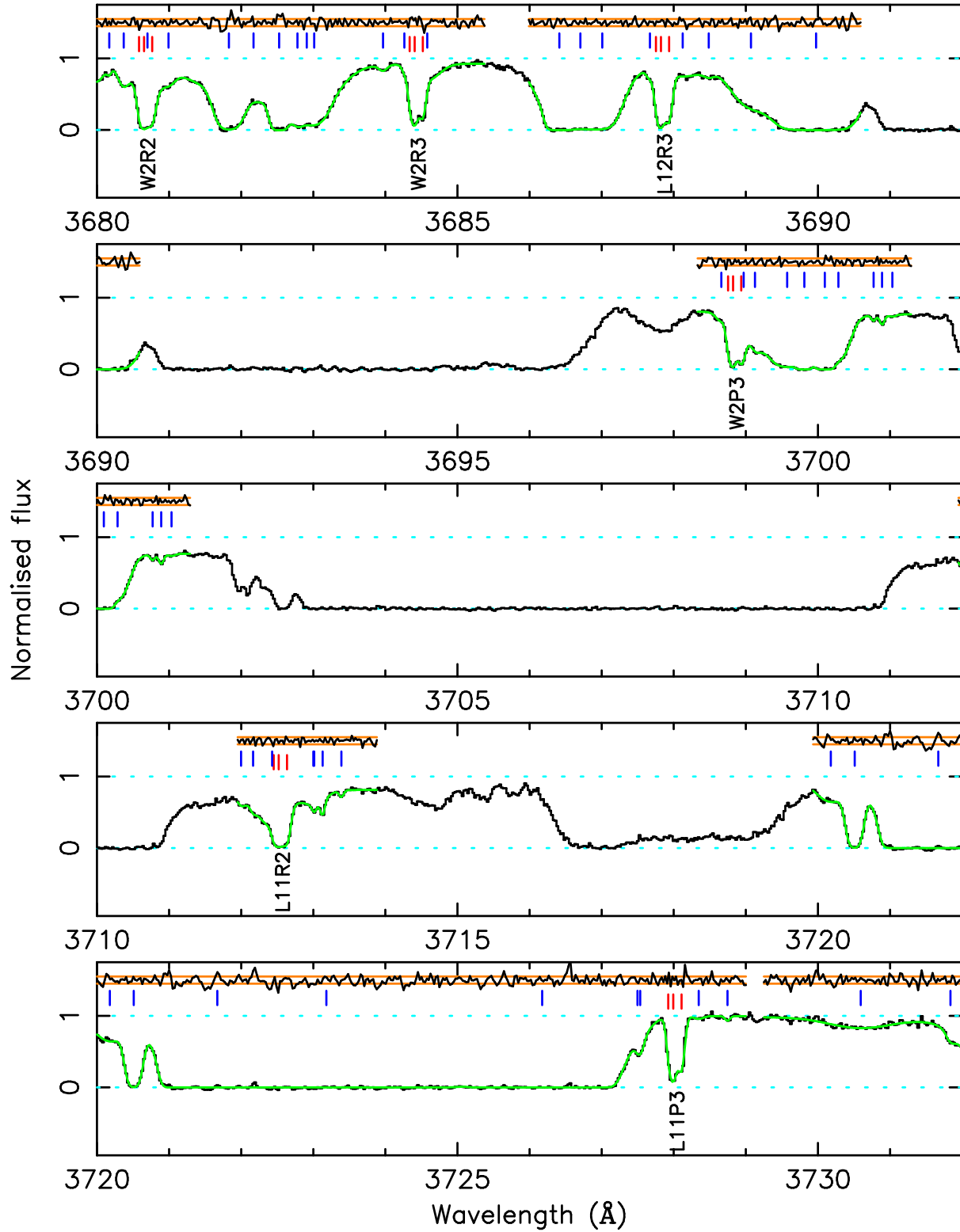


Figure A4. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 4). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

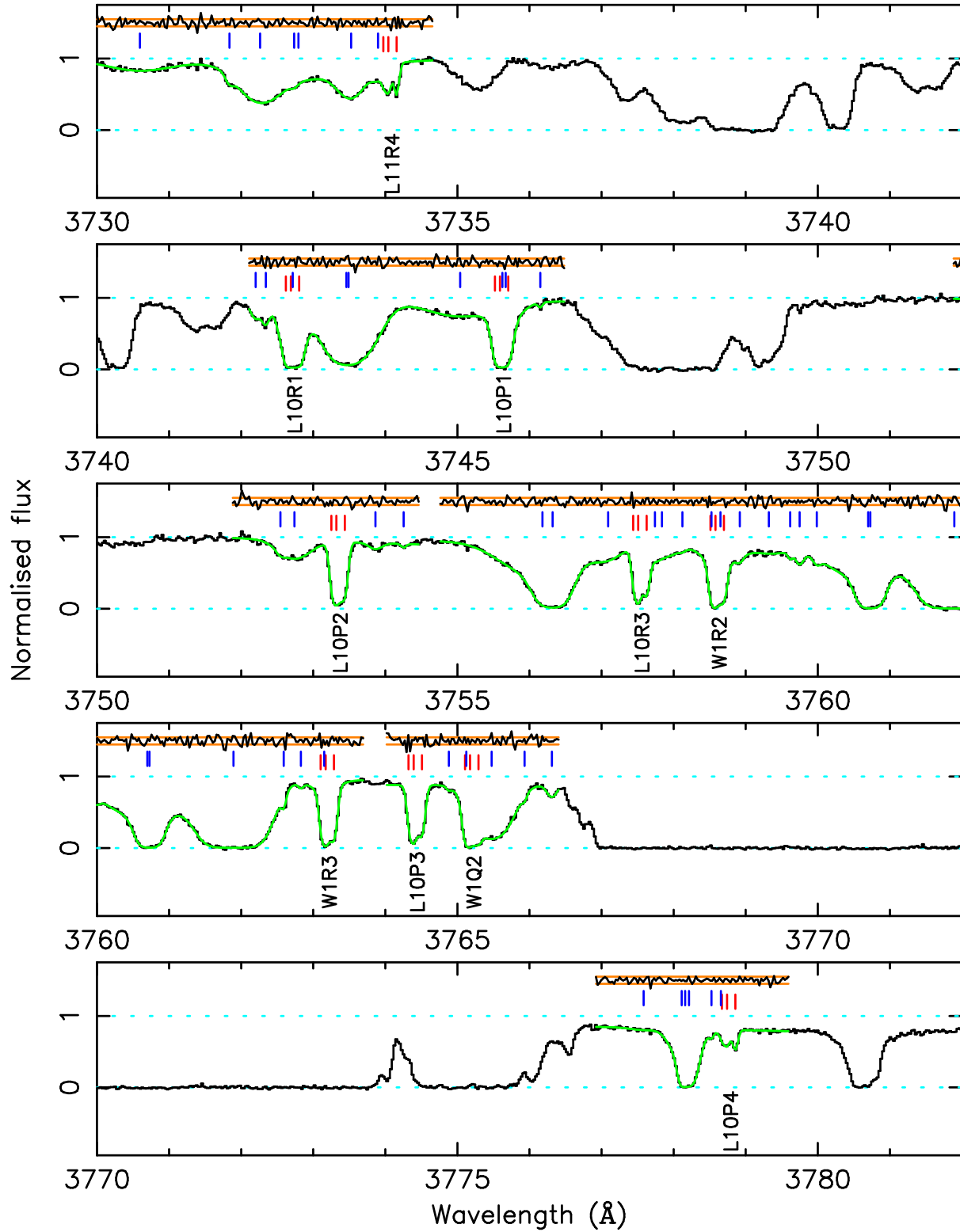


Figure A5. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 5). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

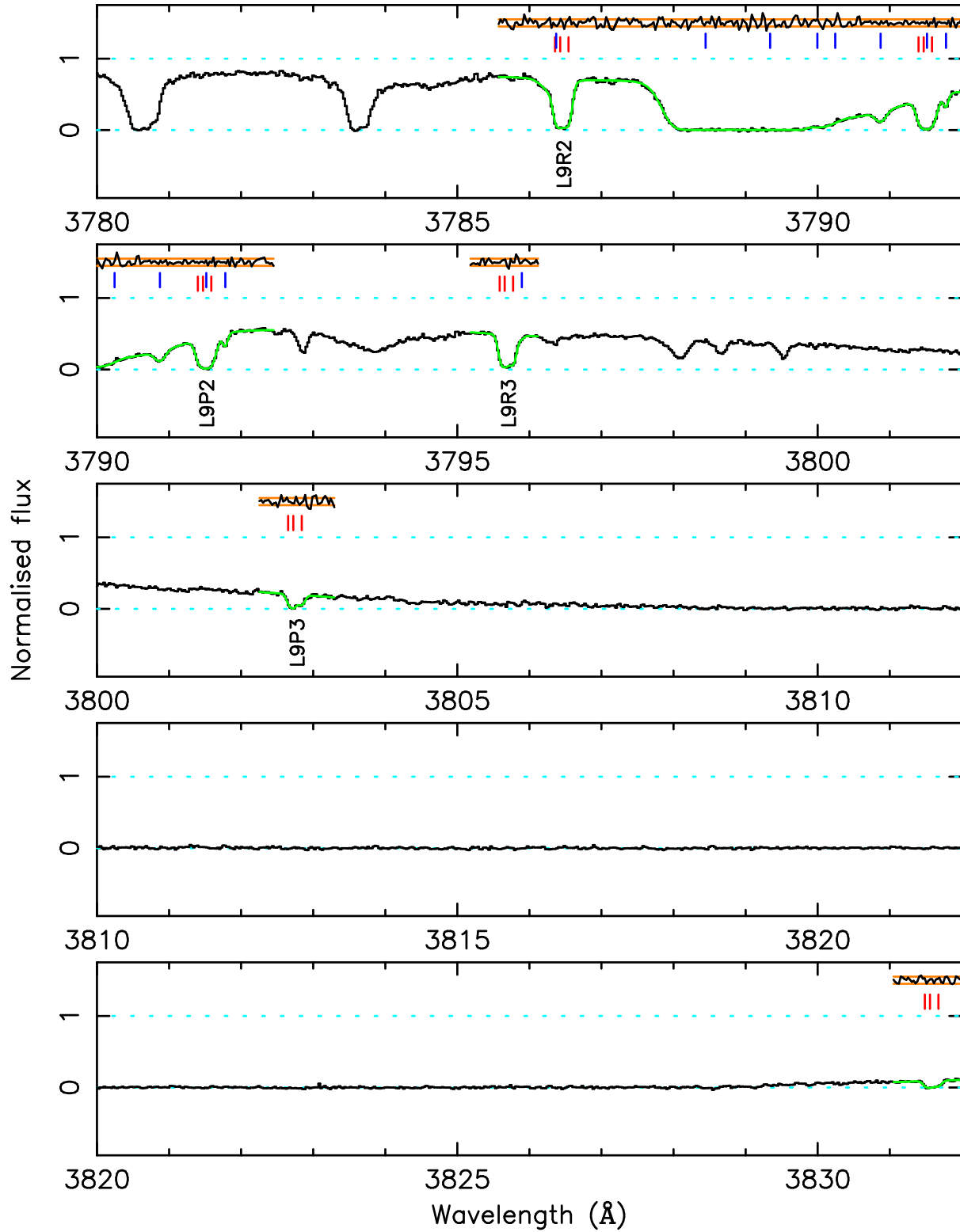


Figure A6. H₂/HD fit for the $z = 2.811$ absorber toward Q0528-250 (part 6). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

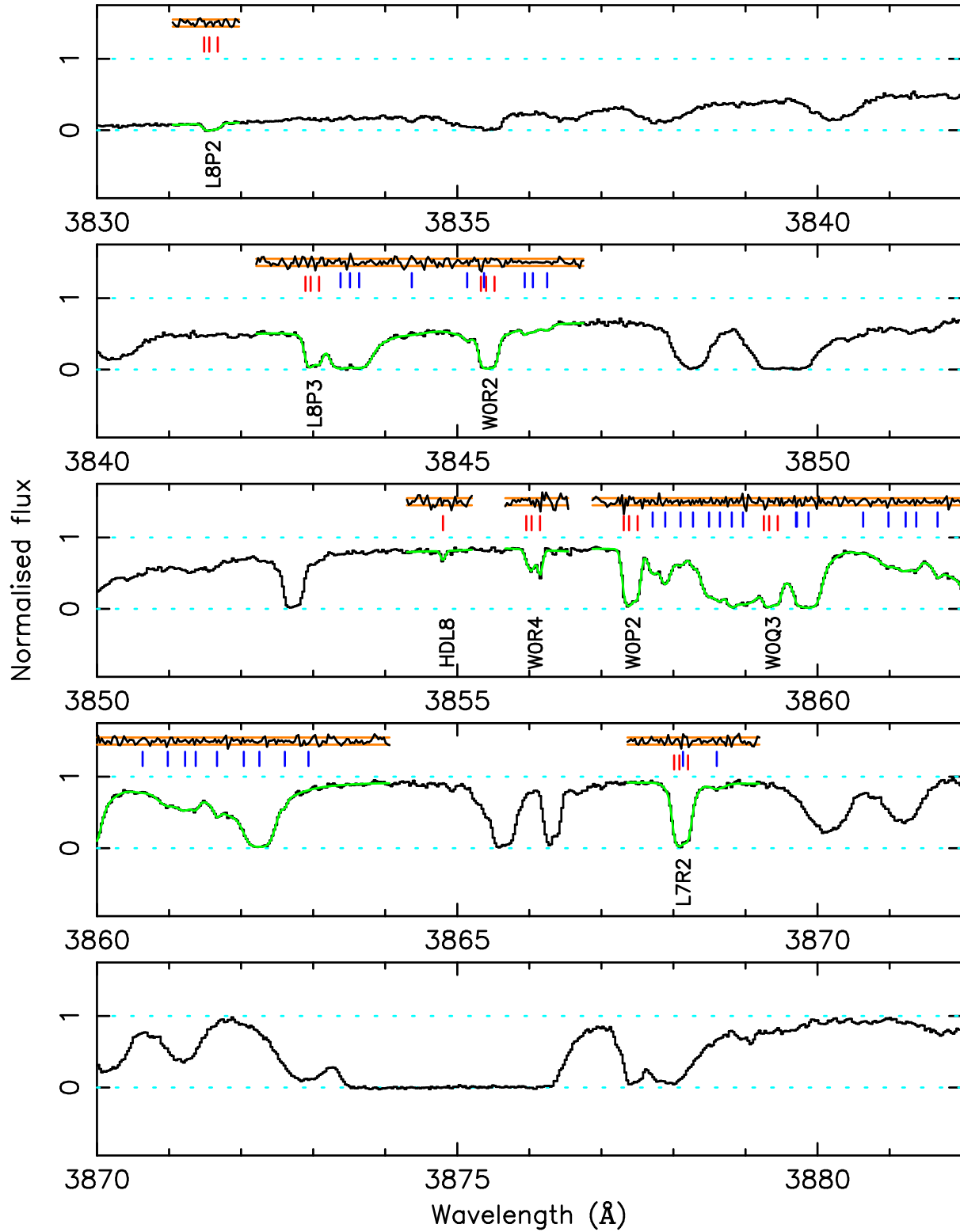


Figure A7. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 7). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

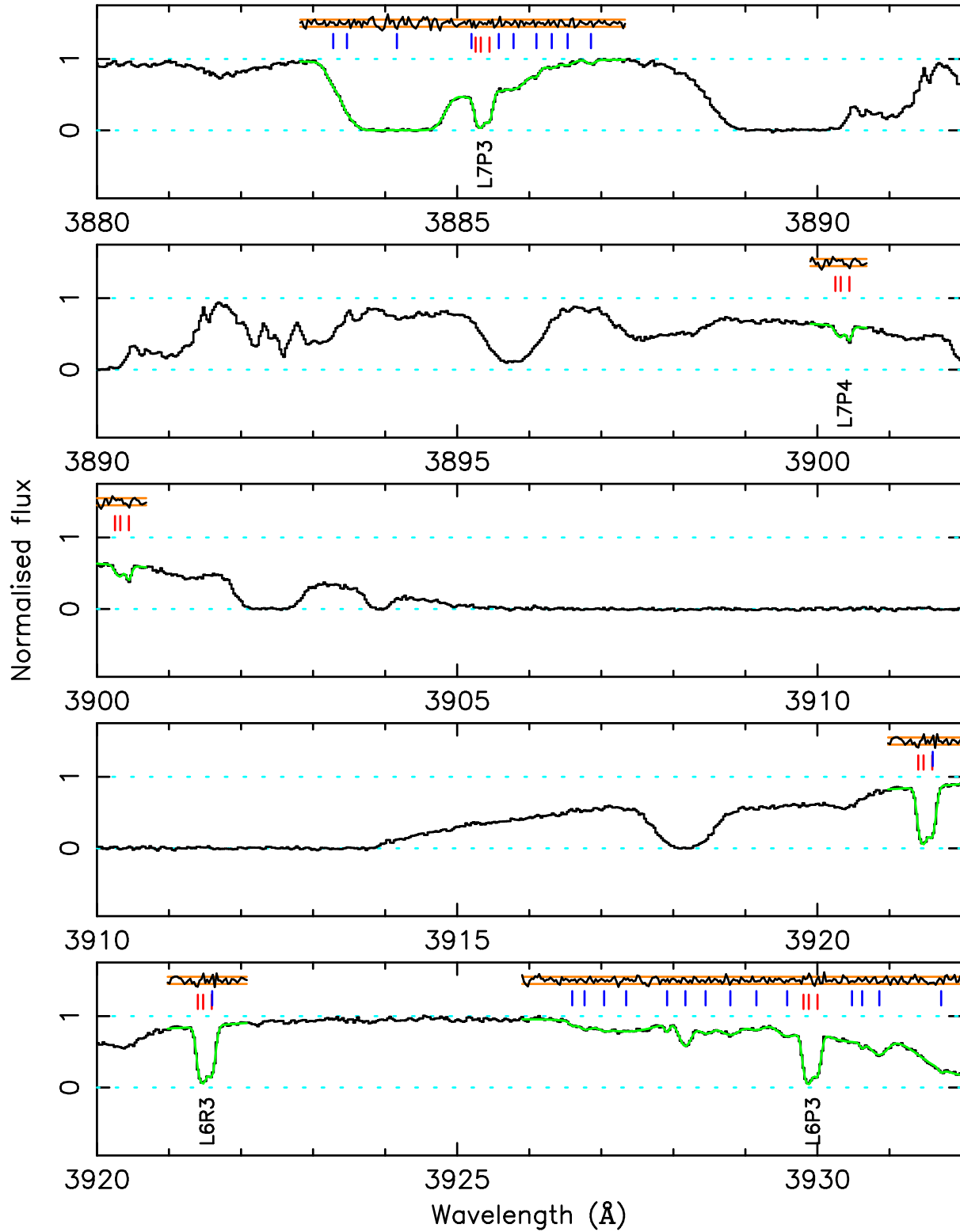


Figure A8. H₂/HD fit for the $z = 2.811$ absorber toward Q0528-250 (part 8). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

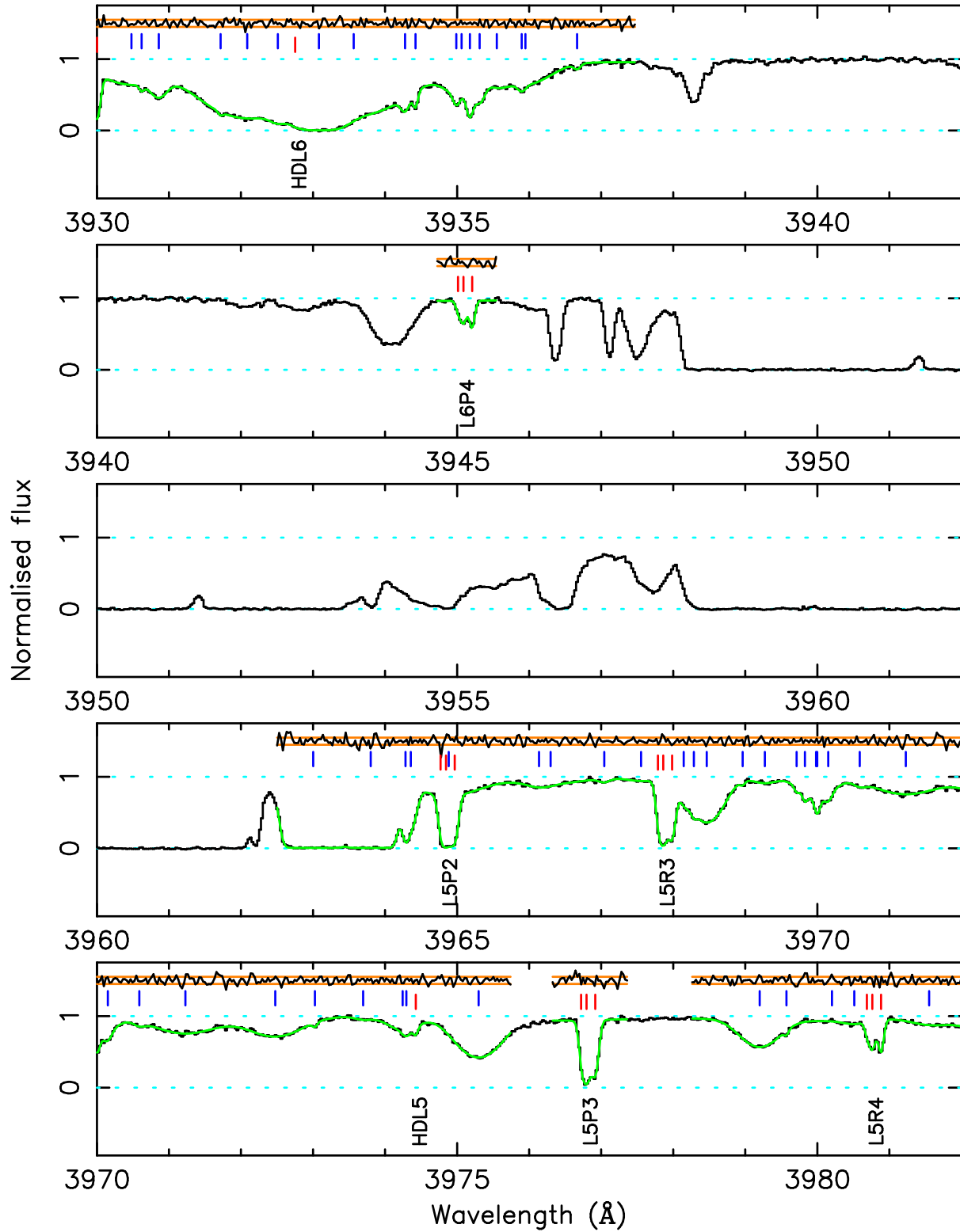


Figure A9. H₂/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 9). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

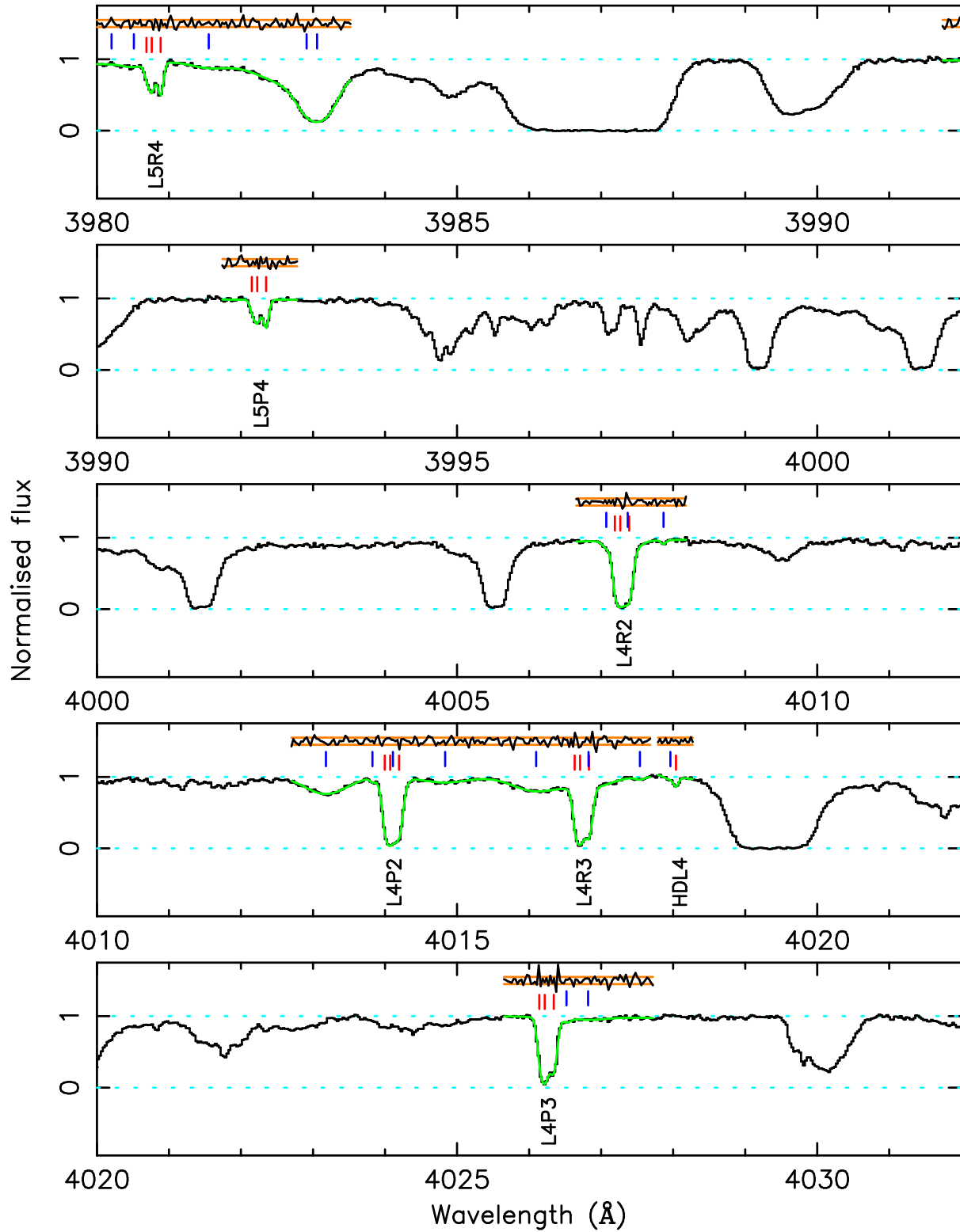


Figure A10. H₂/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 10). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

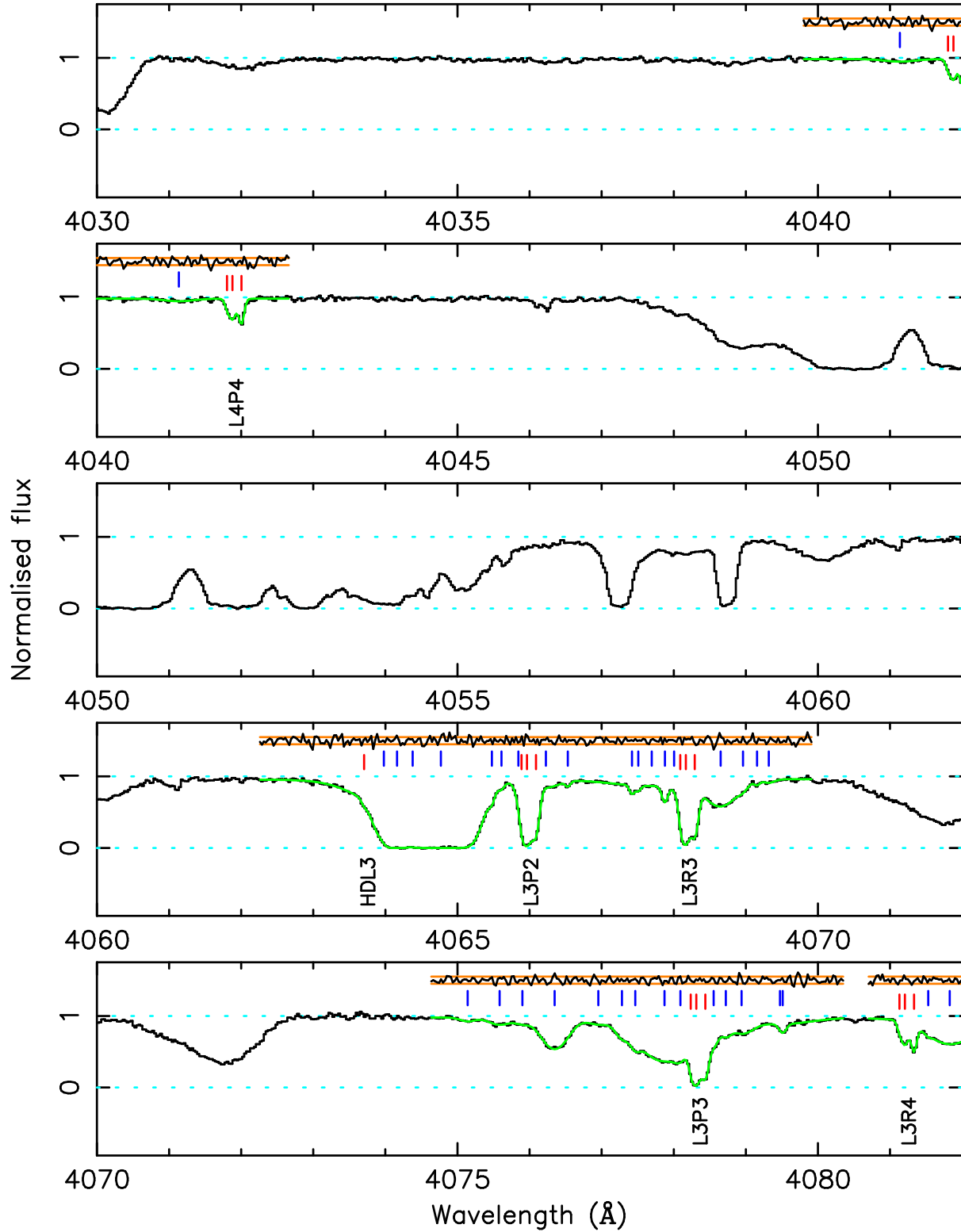


Figure A11. H₂/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 11). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

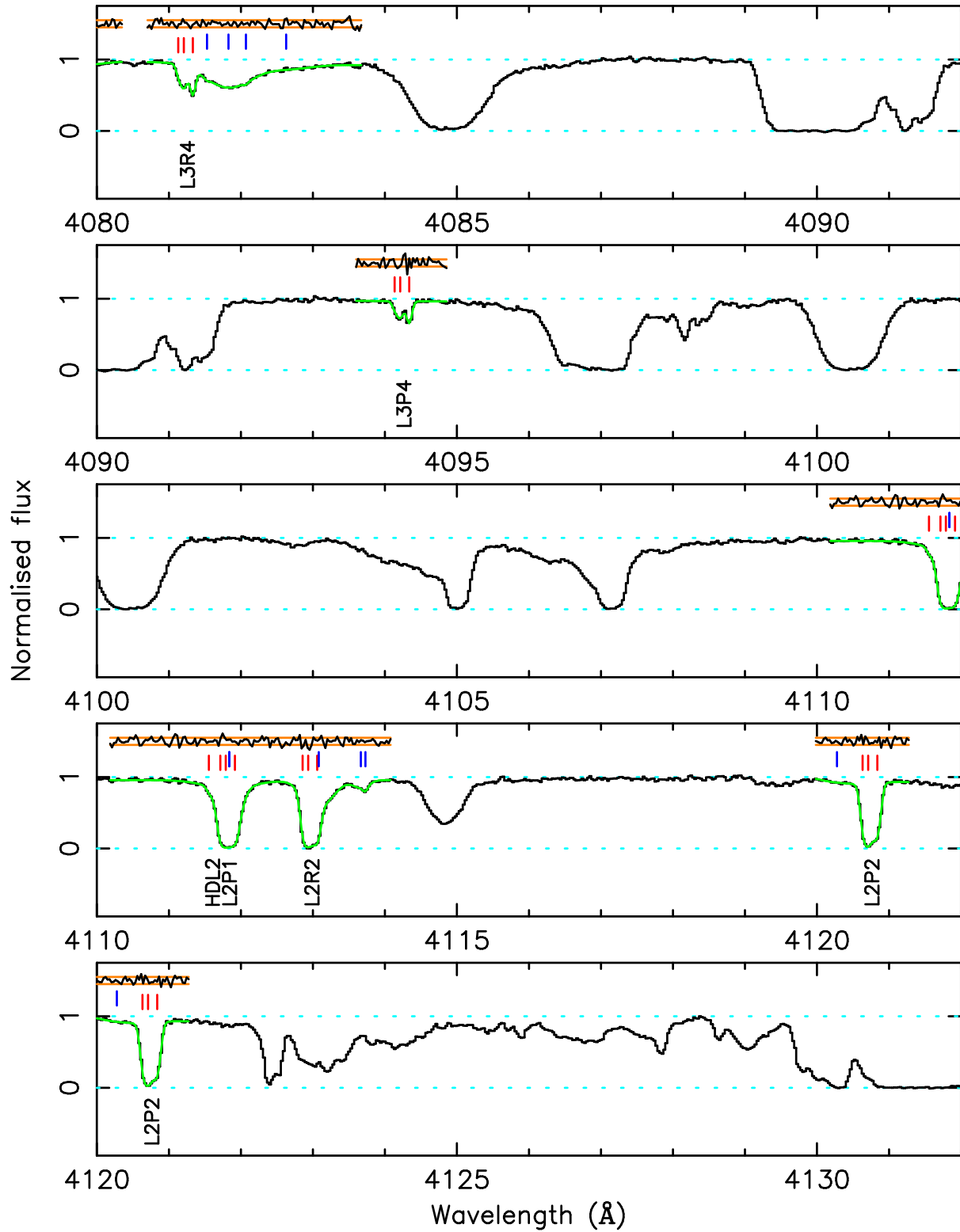


Figure A12. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 12). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.

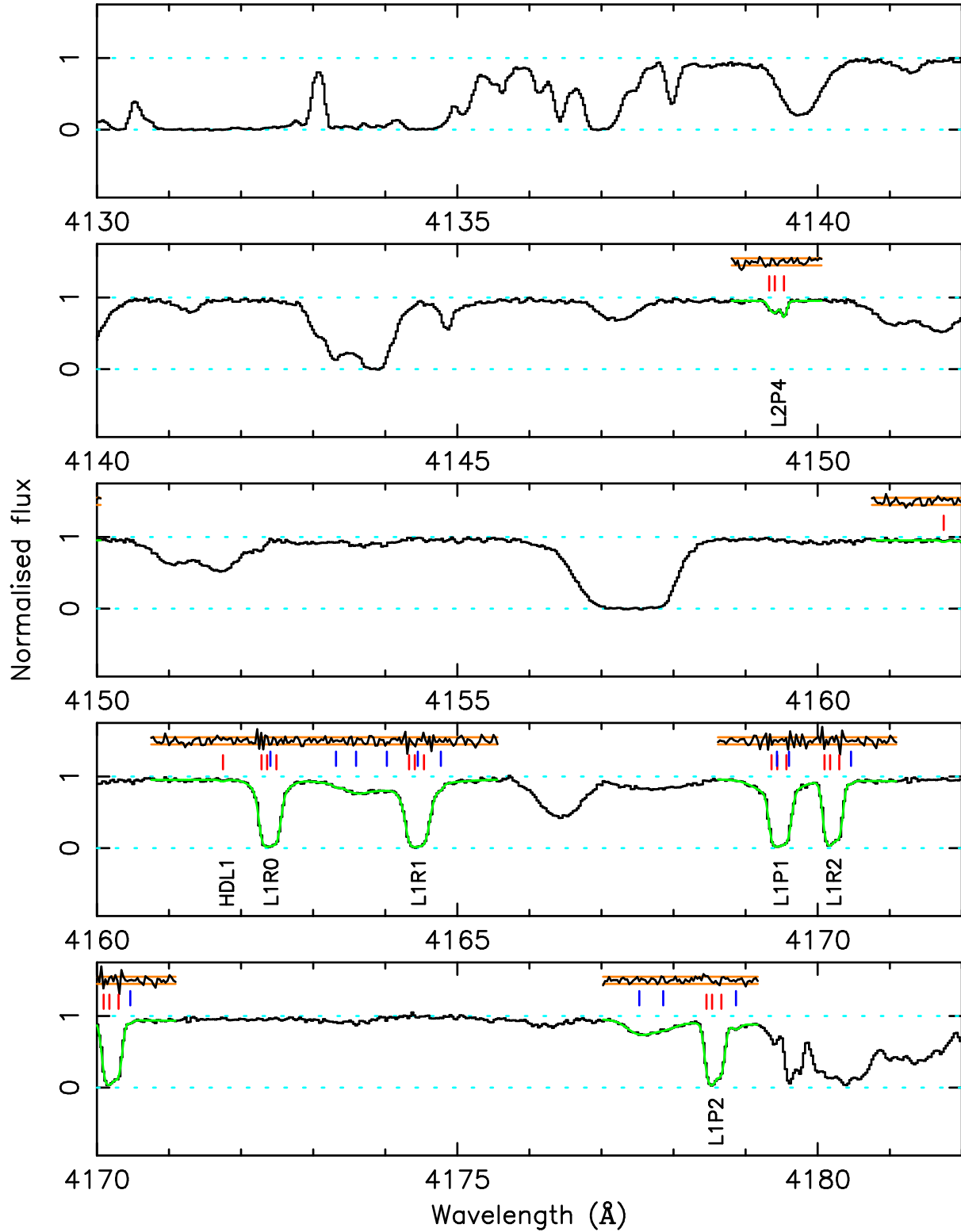


Figure A13. H₂/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 13). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H₂/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. [data - model]/error) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H₂ transitions are plotted below the data.

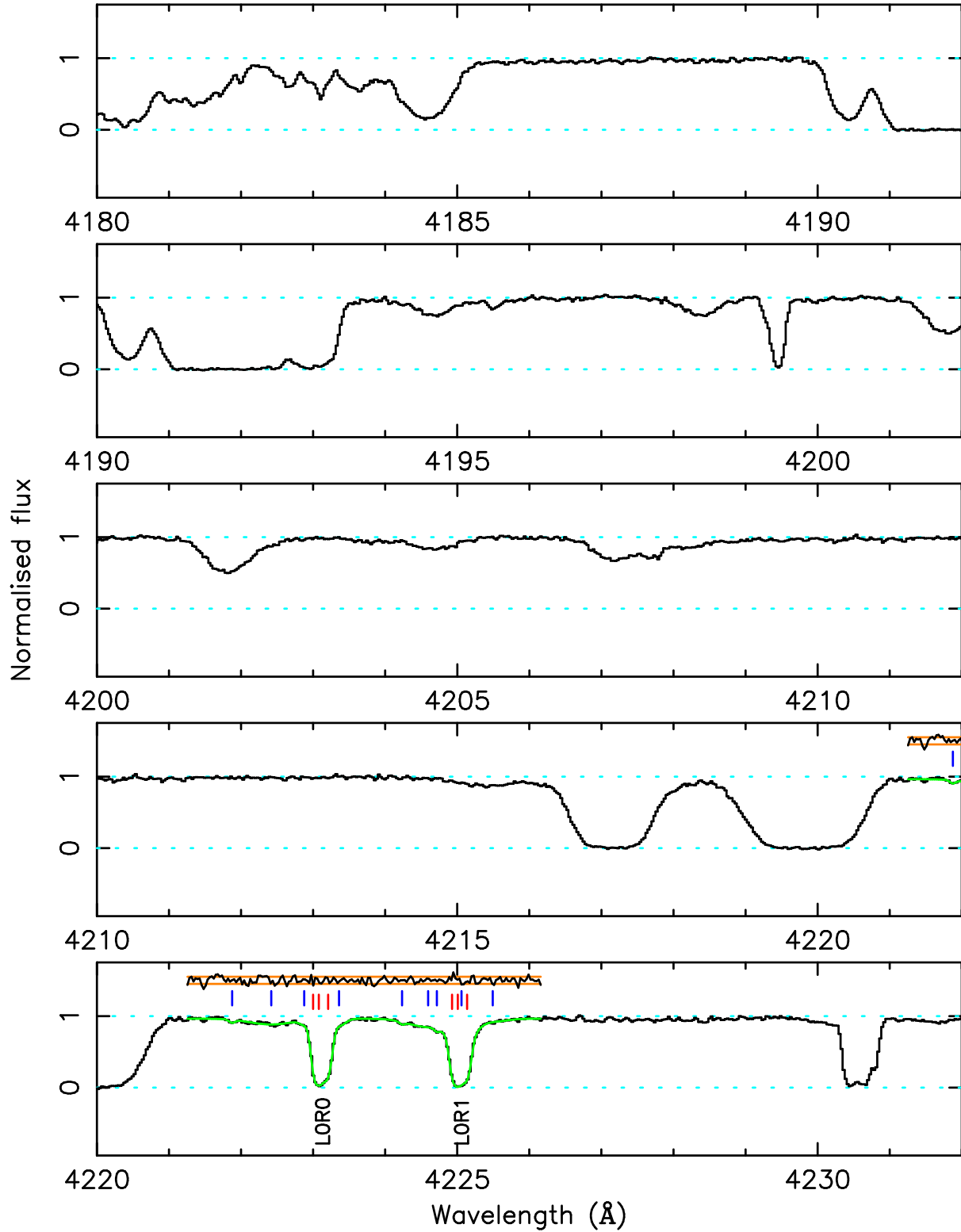


Figure A14. H_2/HD fit for the $z = 2.811$ absorber toward Q0528–250 (part 14). The vertical axis shows normalised flux. The model fitted to the spectrum is shown in green. Red tick marks indicate the position of H_2/HD components, whilst blue tick marks indicate the position of blending transitions (presumed to be Lyman- α). Normalised residuals (i.e. $[\text{data} - \text{model}]/\text{error}$) are plotted above the spectrum between the orange bands, which represent $\pm 1\sigma$. Labels for the H_2 transitions are plotted below the data.